Spring 3-19-2013

Engineering And Pilot Proxy Evaluation Of A Spoken Web Enabled Care Management Platform

Rema Padman
Carnegie Mellon University, rpadman@cmu.edu

Erika Beam
Carnegie Mellon University

Rachel Szewczyk
Carnegie Mellon University

Arun Kumar
IBM, India

ENGINEERING AND PILOT PROXY EVALUATION OF A SPOKEN WEB ENABLED CARE MANAGEMENT PLATFORM

Rema Padman[a], Erika Beam[a], Rachel Szewczyk[a], Arun Kumar[b], Sheetal Agarwal[b], Monia Puri[b], Meghna Singh[b], Jyoti Grover[b], Amit Nanavati[b], Pawan Khera[c]

[a]*The H. John Heinz III College, Carnegie Mellon University, Pittsburgh, PA USA*
[b]*IBM Research Labs, India,* [c]*IBM Research, USA*

Email: rpadman@cmu.edu; beamel.eb@gmail.com; rszewczy@andrew.cmu.edu; kkarun@in.ibm.com; sheetaga@in.ibm.com; moniapur@in.ibm.com; meghna.singh@in.ibm.com; mailto:jyogrove@in.ibm.com; namit@in.ibm.com; pkhera@us.ibm.com

Abstract
Telephones are a ubiquitous and widely accepted technology worldwide. The low ownership cost, simple user interface, intuitive voice-based access and long history contribute to wide-spread use and success of telephones, and more recently, that of mobile phones. This study presents our preliminary efforts to leverage this technology to bridge disparities in the access to and delivery of personalized health and wellness care by developing and evaluating a Spoken Web enabled Care Management Platform (SW-CMP) for underserved and disadvantaged populations. A pilot evaluation study was conducted with eight proxy users representing both providers and patients. Surveys completed by the participants and data generated from the evaluation process were analyzed. Results were mixed; while the current state of the system does not yet fully support a beneficial user experience and needs significant re-engineering, promising opportunities were identified and are being pursued for improved health management for patients and providers alike.

Keywords: Chronic disease management, Health and wellness care, Telephone technology, Usability testing, Proxy evaluation, Underserved populations

1.0 Introduction

Patient-centric solutions are evolving to empower individuals to better manage their own wellness, lifestyle and healthcare needs and potentially serve as a mechanism to help address healthcare quality and cost issues (Hibbard et al., 2013, Cohen et al., 2012, Fisher et al., 2012, Baiker et al., 2010, Bodenheimer et al., 2003). Currently,

most technologies operate on a web-based platform requiring access to the internet or are solely designed for use on computers. Consequently, a critical subset of the population, such as the functionally illiterate, elderly, disabled, and underserved, may not have access to these technologies and the benefits they offer (AHRQ, 2011).

Product and service developments catering to this population specifically are scarce (Sommer, 2009; NCIOM, 2008). If the digital divide is to narrow in the future, there must be a better balance of knowledge dissemination in both the wealthy 10% and the other 90% of the population (Norris, 2000). It has been recognized that over 70% of the world's mobile telephone users live either in an emerging or developing country, thus the telephone is a promising technology to leverage to help bridge the knowledge gap (ITU, 2012).

A novel technology called the Spoken Web (SW) that utilizes a telephone-based network instead of the Internet has been developed by IBM India Research Labs (IBM IRL) (Kumar et al., 2010). Leveraging this technology platform, we have conceptualized and engineered a prototype healthcare delivery and management application targeted at the critical demographic of underserved and disadvantaged populations worldwide. The ultimate goal is to develop a content-rich, Spoken Web enabled Care Management Platform (SW-CMP) that (1) allows stakeholders such as patients, providers, allied healthcare professionals and caregivers to create, update, manage and monitor an individual's health records, (2) provides the capability for social collaboration in managing personal health, (3) enables privacy, confidentiality and security of information, and (4) integrates evidence-based guidelines for lifestyle and behavior modifications and self-health management through educational podcasts, announcements, reminders, collaborative browsing and social media. Given the novel and unique nature of a voice-driven health information and management platform, multiple evaluation studies are needed to establish a deep understanding of the technical, socio-cognitive, domain and logistical challenges involved in the successful adoption and diffusion of such an application. This study reports on our preliminary proxy evaluation of the basic prototype implementation.

1.1 Spoken Web

Spoken Web (SW) is an availability technology that makes voice-controlled interactive information available over narrowband networks using voice-only devices

(Kumar et al., 2007, 2010). It functions analogously to the World Wide Web (WWW) in that it allows navigation between different sites by voice commands. The most notable difference between the WWW and the SW is that while the WWW is delivered in a written format, the information in the SW is stored and delivered in an audio format, thus the content on SW is only available through the telephone network (Agarwal et al., 2009).

The SW employs an easy-to-use voice interface called a VoiceSite over the already in-use phone system to create dynamic content in the local language. End users of the SW can create VoiceSites (or VoiSites) using only the telephone, enabling low literacy users to also create them (Kumar et al., 2010). Users can customize by recording information about themselves and storing it for retrieval at a later time. The VoiceSites can also provide a location for a public forum by using a telephone number that can be called by all members of a user group. Information contained in VoiceSites is maintained as structured content in a database while unstructured content such as social media discussions in the public forum are stored as XML files (Agarwal et al., 2009, Kumar et al., 2007, 2010).

The VoiceSites are generated by an application, VoiGen, that configures voice applications to enable a caller to easily access VoiceSites. This is similar to how people can access customized web pages. After it has been created by VoiGen, the VoiceSite is hosted on the network by VoiHost. VoiHost is an engine that resides in the telecommunication provider's infrastructure and can be shared by many subscribers (Kumar et al., 2007).

Analogous to a website, a VoiceSite can give the user an opportunity to navigate to other VoiceSites through links that utilize only audio (see Figure 1). The connections are made using Hyperspeech Transfer Protocol (HSTP), which is analogous to Hypertext Transfer Protocol (HTTP) used by WWW (Kumar et al., 2010). When a user wants to move from one VoiceSite to another, s/he can speak into the phone as he would a normal call and be transferred to the next VoiceSite of his choice. These links are pre-defined in the system and can be navigated easily without the need to take the phone away from the ear. This “browsing-by-talking” experience is useful, and fosters a similar navigation capability as “back buttons and bookmarks” in the World Wide Web (Kumar et al., 2009).

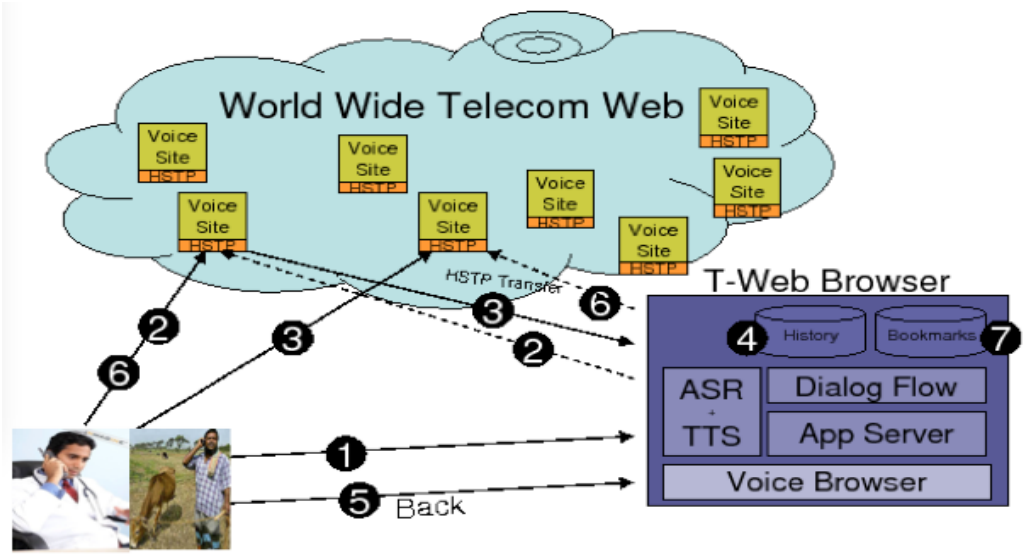

Figure 1. Browser for Spoken Web (Agarwal et al., 2009)

1.2 Chronic condition management

Spoken Web is a versatile platform that can be leveraged for use in a variety of industries. In the healthcare arena, it can be particularly useful for supporting care for patients with chronic conditions as it has the potential to help patients self-manage their care. According to the World Health Organization, chronic diseases are the leading cause of more than 60% of deaths around the world, more than that caused by communicable diseases (WHO, 2013). About half of Americans today, 149 million, are affected by at least one chronic condition and about 25% of Americans have two or more (ADA, 2012). Patients with chronic conditions impose a large burden on the health system due to the continuing cost of their care; finding a way to reduce these costs is paramount. Both developed and developing countries are becoming increasingly affected by chronic diseases; chronic diseases account for 70% of health care expenditures and are the main reason for physician office visits (Holman et al., 2004, Sokol et al., 2005).

The large number of people who experience chronic conditions face unique challenges including ongoing experience of symptoms, lifestyle changes, obtaining useful medical attention, and managing their medications, among others (Hall et al., 2003). Collaboration between patients and medical professionals, as well as family members of the patients, can lead to improved outcomes (Wicks et al., 2010). Patients can also do a few things to better manage their condition; they can develop

skills to improve self-efficacy, develop greater self confidence, surround themselves with a supportive environment in family, workplace, and health care environments, and effectively monitor and react to changes in their conditions or the associated symptoms (Cohen et al., 2012). Applications like the SW-CMP can potentially provide patients with chronic conditions with a means to gain these skills, and tools to help manage their own care.

Given the objective of supporting patients with chronic conditions using SW-CMP, we have developed a prototype that can potentially help patients with diabetes manage their health more efficiently and enhance the physician-patient interaction to support better care management. Globally, about 285 million or eight percent of the world's population was affected by diabetes in 2010 and it is expected to double by 2030 (IDF Diabetes Atlas, 2012). About 80% of individuals with diabetes live in low and middle-income countries, and it is expected that developing countries will carry more of the burden in the future (Diabetes Facts, 2012). This disease is the cause of many deaths worldwide. The disease also has an astronomical economic impact, with $471 billion spent globally on diabetes related health care by the end of 2012 (IDF Diabetes Atlas, 2012). This is largely due to the complications associated with diabetes when it is not effectively managed.

Patients with diabetes, on average, have care costs 2.3 times as much as those without diabetes (Diabetes Statistics, 2012). With the growing number of people affected, it is becoming increasingly important to monitor and manage the treatment of diabetes. Disease management programs have many benefits including health benefits to patients and economic benefits to patients, providers and payers (Gray et al., 2012). Patients involved in disease management programs have less hospital visits and shorter lengths of stay (Cohen et al., 2012). It has also been shown that there is a decrease in the number of visits to outpatient care settings when the management system is in place. There is an important shift that must be made from a focus on episodes of care to the whole care process, from the start of a chronic disease treatment plan through the entirety of a patient's life.

Although diabetes management is imperative in improving patient outcomes, it is not easy to establish a seamless and effective management system; there are many obstacles that must be overcome. Obstacles include information sharing and seamless communication between and among patients and providers (Mirza, 2008). Additionally, behavior change for newly diagnosed diabetes patients is a challenge, as

is maintenance of healthy behavior once it has been first achieved (Hall et al., 2003). Many have found that there needs to be a support structure in place in order to facilitate the maintenance of a healthy lifestyle (Wicks et al., 2010, Hall et al., 2003). There must also be an effective means of connecting patients to social structures to ensure ongoing success. This has traditionally been accomplished through face-to-face contact between physicians and patients, and has more recently been expanded to web-based systems; however, many patients with diabetes may not have access to the Internet, and therefore cannot benefit from such applications. It also may be hard to maintain a focus on care management if face-to-face contact, with its limitations on time constraints, is the sole means of communication between the patient and team of care providers (Hibbard et al., 2013).

Another option is a telephone-based system, which has multiple advantages including portability, convenience, customizability, and immediate connectivity (Mirza, 2008). One of the most important advantages is its ability to reach people who do not have access to the Internet, but do have access to the telephone. Many people in developing nations own a telephone and can make and receive calls at low cost (WHO, 2011), thus SW-CMP is potentially a viable mobile health alternative for this population.

In the next section, we describe the key features of the Spoken Web enabled Care Management Platform (SW-CMP) designed to support these chronic disease management requirements.

2. Components of SW-CMP

SW-CMP provides an alternate information and communication platform for supporting healthcare delivery and management that has the potential to be particularly advantageous to those individuals who may otherwise not have an opportunity to utilize the information and tools provided via the Internet. SW-CMP includes multiple features that could support the diabetic patient to help manage his/her condition. These features can support patient improvement in many areas: uninterrupted medication use, lifestyle behavior change, effective communication with medical professionals, and better education. The features are also meant to help empower patients and increase self-efficacy (Mizra, 2008). The following features of

the SW-CMP have been conceptualized to support one or more of these areas and add value to the care delivery process for both patients and health care providers.

2.1 Medical profile

The Medical Profile component facilitates collaboration between the patient and medical professionals through shared awareness of the patient's current and past medical information. This feature enables medical professionals to access patient information and update their records to include the latest measurements of blood pressure, height, weight, blood sugar, severity level of diabetes, as well as additional comments that the clinician may want to record. The system has the capability to transfer the current medical information to the medical history section when a newer record replaces it; thus, one action simultaneously updates the current medical profile and the medical history component. Patients can also access this information at any time, empowering them to track progress with their diabetes condition and its management.

2.2 Announcements

The Announcements component of SW-CMP allows providers to record announcements on any number of subjects relevant to the population being served. Announcements may include clinic hours, notification about new medications, and health tips. Patients can listen to the announcements and navigate their way through the full set to those that are most interesting to them using specialized commands, including 'repeat' and 'skip'.

2.3 Queries

The queries component enhances the interaction between the patient and provider by facilitating a means of direct and secure communication. Patients can post sensitive queries that only providers can access and respond to. More general queries can also be posted to the broader community of users. Other patients and providers can listen to these general posts, which may maximize the social networking and social support benefits of the system for users (Fisher et al., 2012). This feature has the added value

of highlighting the similarities and differences between patients and their treatments, and leveraging the concept of social media to improve disease management. Patients may feel comforted to know that other people have the same problems they do, and thus may feel more confident about their own ability to manage the disease if other people can successfully use the system to manage their condition (Wicks et al., 2010).

2.4 Educational Podcasts

It has been shown in the past that users highly value the education portion of a disease management system (Goldman, 2008). The education component offers podcasts about many aspects of the disease that the system addresses. In the case of diabetes management, the education provided includes general knowledge about diabetes, how to improve the physical activity and eating habits to manage the disease, and the effects of the disease on various parts of the body. It can be expanded to include other information that may be helpful to the users and can be stratified in order to group podcasts with similar subjects together to facilitate easier navigation.

2.5 Privacy, Security and Confidentiality

A user accesses the application only through a pre-registered number, and a four-digit password. This provides two-levels of security. One needs to have a phone and know the password in order to access the content. From a confidentiality point of view, no other user will be able to listen to data about another user, unless the user wants them to. Further, Spoken Web supports “roles” and “privileges” such that based on the role, users may be: (a) unable to access a piece of information, (b) only listen to it (read only), or (c) modify it (listen-record).

The combination of these five main components give patients and providers the ability to record and access information that they need to better manage the patient’s chronic condition. These components can be combined in an intuitive flow, which is depicted in Figure 2.

2.6 Navigation Paths

A user's navigation between any two components of the application can be tracked as shown in Figure 2. For instance, a user accesses the system by calling the VoiceSite number and starts in the leftmost component, 'call initiation'. After he logs in, he listens to a welcome message and is directed to a main menu where he hears the options: medical profile, announcements, queries, and education; from there, he can advance to any option that he chooses. Back options are available at each stage so the user can navigate back to the main menu and access a different portion of the system. A user can also ask to repeat an option, skip or exit it.

3.0 Usability Evaluation of SW-CMP

Human factors driving usability issues in technology design have long been studied in the human-computer interaction literature (Nielsen, 1994). Wharton and Lewis (1994) examined the use of psychological theories in developing usability concepts, such as perceptions based on hearing vs. sight, speed and response time, breadth and depth of user's knowledge, role of memory where 'recognition is easier than recall', and finally, user's intention with regard to the technology. While these concepts have been tested in web-based and other computer-based technology design settings, its application in the context of voice-based healthcare delivery technologies is a new challenge that we aim to address in this research. Limited studies have explored the use of voice-based communications, such as for data collection in low resource settings (Lerer et al., 2010), identifying its advantages as lower cost, better accuracy, and effectiveness in collecting sensitive information compared to a live interviewer (Patnaik et al., 2009, Tourangeau et al., 2003). Some challenges associated with voice interfaces have also been documented, including the difficulties in following local conventions, limited abilities of the human mind in processing information, user preferences and differences, and current limitations of speech recognition technologies (Suhm, 2008). These challenges have resulted in many studies reporting low success with evaluations of voice based applications in the field.

Prior to evaluating SW-CMP with actual users to understand its impact on health, financial and organizational outcomes, it was thus important to assess the basic

usability of the application and subsequently re-engineer components that failed to meet end user needs. As indicated earlier, usability of a technology refers to the ease, efficiency and satisfaction in using or learning to use the technology by actual end users (Nielsen, 1994). However, given the nature of the target population for whom SW-CMP is intended, the challenges in conducting the early testing reported in this study led to the use of proxies. Furthermore, while benefits from using SW-CMP are anticipated to accrue for the target patient population defined earlier, education, management and monitoring are expected to be provided by allied health professionals such as community health navigators. Hence it is equally important to evaluate the usability of SW-CMP with these secondary users. Proxies have been successfully used in prior studies in the literature in speeding up design iterations of technologies and in medical care (Sy, 2007, Pickard and Knight, 2005), hence we employ this approach to conduct our preliminary evaluation of SW-CMP before testing it with actual users from the intended target population.

Thus, a pilot evaluation study was conducted with proxy users in order to assess the application for usability, clarity and effectiveness of communication, accessibility, and ease of navigation. This proxy evaluation approach required study participants to play the roles of actual end users including patients and providers. The study also utilized a survey that was developed specifically for an evaluation of SW-CMP users. Using this survey data, along with click-stream data from system log files and structured information captured in the application database during the evaluation, results were analyzed to understand participant opinions of the SW-CMP, the usability of the technology, and their ability to take advantage of the SW-CMP's features to improve health management.

3.1 Development of Evaluation Survey

The survey was designed to gather participant opinions of the usability of the features, the voice interface, and the potential for the system to accomplish tasks that would be done by typical end-users; the survey was thus a blend of a number of validated questions from multiple surveys, with an underlying structure modeled by the US National Institute of Standards and Technology (NIST) guide for improving usability of electronic health records (EHR) (Schumacher et al., 2010).

The interface survey questions were adapted from the Questionnaire for User Interface Satisfaction (QUIS^{TM}). These survey questions are categorized into sections focused on overall reaction, terminology and system information, learning, and system capabilities (QUIS^{TM}). Originally designed for use with computational interfaces, relevant questions were selected and adapted based on their applicability to the SW-CMP. The usefulness of the system was evaluated using survey questions derived from Perceived Usefulness and Ease of Use survey instrument (Davis, 1989). Consistency measures were derived from (Nielsen, 1994). Other user satisfaction questions were derived from the After-Scenario Questionnaire (Lewis, 1995). All questions were adapted as necessary to fit within the given study context.

3.2 Usability Test Preparation

User profiles were created in the system prior to testing. This included recording information such as name, birth date, phone number, gender, and role (patient or provider) such that participants could assume the role of an existing user. Additional information about patients was also recorded, such as health data, so that the evaluator could effectively use the medical history and medical profile sections (drawing on the most recent and past medical information which would normally be entered by a provider). After the user profile was created in SW-CMP, the system generated a user ID and password for each participant, which was recorded for later use.

3.3 Usability Test

The usability test of SW-CMP was conducted in Fall 2011. This test included eight participants comprising graduate students and staff at the researchers' university who self-selected to participate in the study. These were educated users who were familiar with phone-based technologies such as caller hotlines and bank services. While this was a significantly different segment of the user population than the target population for whom this technology is intended, this is the user population that will provide training, education and support for the target population, hence it is important to assess its value for such users. The intent of this experiment was to obtain some insights into the usability, navigation, language interpretation and access issues

associated with a new health management platform. None of the evaluators had previous exposure to SW-CMP, nor were they paid for their participation.

A brief overview of Spoken Web and the purpose of SW-CMP was provided at the beginning of the evaluation session. Each evaluator was then given an instruction packet that included a previously-created user profile. The tasks simulated those that a real user in that specific role (patient or provider) would normally want to accomplish using the SW-CMP. Provider tasks included recording patient information, answering queries, and listening to announcements, and educational podcasts. Patient tasks included listening to medical profile and medical history, posting queries, listening to query answers, listening to announcements, and listening to educational podcasts. Each of the participants was asked to record the amount of time spent on each task as well as their level of success in completing the assigned tasks.

The evaluators called the SW-CMP number simultaneously, which allowed them to access the first VoiceSite. They were prompted to punch in on the phone keyboard their username and password, which were provided in their instruction packets. They then moved through the system in the appropriate sequence required by their assigned tasks.

Participants were asked to complete the evaluation survey at the end of the testing session. Questions were based on Likert scales with ranges from 0 to 9 or from 1 to 7. Questions assessing the consistency, clarity, straightforwardness, and ease of mistake correction were scaled from 0 to 9. Questions assessing logic of the system, ease of use, and success in fostering task completion were scaled from 1 to 7. The survey also included open-ended questions asking evaluators to comment on the best and worst features of the system as well as features they thought would be useful to include in future versions of SW-CMP. After the tasks and the surveys were completed, the instruction packets and the evaluations were collected and analyzed.

There were a few problems that arose during the course of the testing. Some users had issues accessing the system; others had issues getting the system to recognize their voice commands and may have become frustrated by the process. They were advised to try the command at least three times; the system forced the user out after three failed attempts. If they were not successful, they were asked to repeat the task and use the dual-tone multi-frequency (DTMF) option in order to enter commands. If they repeated this process three times with no success, they were advised to move to the next task.

3.4 Process Modeling

The modeling of the process that the user goes through in the SW-CMP provides insight into the functionality and usability of the system, and to recognize the features that are most often used and valued, judged by the amount of time used and frequency of access. This unique visual representation of the user experience allows objective measures that complement self-reported user satisfaction. These process flow diagrams are developed using data from the call logs and click-stream data generated by SW-CMP during every call.

Figure 3 illustrates an example process diagram depicting the flow of a call and the steps a caller went through while interacting with the system. The arrows at the top of the diagram denote components that the user accessed. For sections where the user was in multi-level components (e.g., choosing an option from the main menu and entering the selected part of the application), a second arrow was placed under the first in order to show the component hierarchy. Within each component, individual procedures were represented as two or three character codes. The start and end times for each process are also shown in the figure. The bottom arrow shows the direction from oldest action to newest, going from left to right. If there was an irregularity within a call, a callout box was used to note what it was and where it occurred during the call process.

4.0 Results

4.1 Analysis of Navigation and Processes

Click-stream data from system log files generated during the evaluation process were analyzed to model the process diagrams discussed earlier. These process diagrams allowed insights into users' navigation through the system.

Overall, there were 55 calls from the eight proxy users in the study; the number of calls per user ranged from four to nine (median of 6.5). Using the process diagrams, the average time the user spent in each component and the total time of the call were calculated. On average, evaluators in a successful call spent 2:03.92 minutes in

announcements, 6:21.58 in education, 0:48.59 to submit a query, 3:35.47 to listen to a query, and 1:00.83 to answer a query. None of the evaluators successfully completed the profile section, and therefore an average time could not be calculated. The average length of a call with and without errors were also calculated. An evaluator who did not encounter an error spent an average of 5:24:26 minutes on the call while one who did encounter an error spent an average of 1:01.90 minutes; overall, all users spent an average of 2:03.91 minutes.

Using the diagrams, many calls with potential issues were identified. There were three types of issues found: the application closing unexpectedly (51%), the SW-CMP encountering a program error (13%), and the improper closing of a component (11%). Out of the 55 calls, 29 calls (64.4%) had occurrences of such instances, with some calls having multiple types of errors.

4.2 Analysis of Usability Survey

All twenty questions on the evaluation survey were answered by at least seven evaluators, and most were answered by all eight evaluators. Results are displayed in Table 1. Evaluator responses to survey questions ranged across the full spectrum from very dissatisfied to very satisfied. The actual meanings of dissatisfied and satisfied responses were different based on the context of each question, and are labeled more generally for reporting purposes. The questions and the true response options are listed in Table 1. In all scales, low numbers indicated dissatisfied responses and high numbers indicated satisfied responses. Dissatisfied responses include those that ranged from 0 to 3 for the questions with a 0-9 scale and from 1 to 2 for the questions with a 1-7 scale. The survey responses were broken down using three groups: all evaluators, provider evaluators, and patient evaluators. This allows for better insight into the general evaluations of the system as well as role-specific evaluations.

Evaluators were satisfied with many portions of the system. They thought the system was very consistent (100.0%) and performing tasks was straightforward (57.1%). Additionally, they thought the ordering of the menu options was logical (37.5%) and the system was designed for users of all levels (28.6%).

At the same time, evaluators were generally dissatisfied with the ability of the user to correct mistakes (43.9%). They also thought it was unlikely that they could get the system to do what they wanted it to do (50.0%) or that they could use the system

successfully every time (50%). They also thought that the current version of SW-CMP needed additional enhancements to meet their needs as users (42.9%).

Many responses fell in the middle range, indicating neutral responses. These included responses from 4 to 6 for the questions with a 0-9 scale and from 3 to 5 for the questions with a 1-7 scale. Just over half of the questions (11) elicited a less-than-favorable response from the evaluators. These included most questions that addressed the system value such as whether the system was terrible or wonderful, difficult or easy, and frustrating or satisfying.

The provider evaluators indicated overwhelming satisfaction with eight of the measures; the strongest indications included that the system was consistent (100.0%), performing tasks was straightforward (100.0%), and the ordering of menus was logical (75.0%). Half of these users indicated satisfaction with clear prompts for input, ease of learning the system, usefulness of the system, likelihood to recommend the system, and confidence that the system would make it easier to track health information. On the negative side, the provider evaluators indicated dissatisfaction with three main points: correcting mistakes was difficult (66.7%); it would be unlikely that they could get the system to do what they wanted it to (50.0%); the system is not user-friendly (50.0%). At least 50% of evaluators responded with neutral feedback for thirteen of the questions, including the questions assessing overall satisfaction with the system (75%).

The patient evaluators responded with more dissatisfaction than the provider evaluators. At least 50% of these respondents indicated dissatisfaction on several questions. Four questions had a 75% dissatisfaction rate: these included indications that these evaluators were not confident that they could operate the system successfully every time. They did not think it was likely they would recommend the system, that it would improve their health, or that it would help keep track of health information. Only one item, confirming consistency of the system, elicited entirely positive responses (100.0%). Twelve questions had at least 50% neutral responses. Six of these questions elicited neutral responses for the provider evaluators; these were questions about ability to correct mistakes, logic of menu options, usefulness of the system, user-friendliness of the system, and overall aspects of the system.

4.3 Analysis of Comments

There were three types of open-ended responses that were collected during the usability testing: the positive aspects of SW-CMP, the negative aspects, and features that would be useful to add to the application. The majority of the positive comments related to the information that the system was able to provide such as the podcasts. Users were also generally impressed with the theoretical potential of the tool. The most negative comments included those based on technical difficulties such as the system understanding the user's voice and the fact that there were no dual-tone multi frequency (DTMF) or Touch-Tone choices listed before the menu options; furthermore, the users commented on the fact that the system dropped calls.

On the clinical side, the participants recommended use of indicators for blood sugar level which would enable diabetes patients to incorporate clinical information into their lifestyle. They also recommended new features to be included that listed do's and don'ts for diabetic patients and additional information about diabetes. It was suggested repeatedly that the application have numbered choices so that it would be easier to use the DTMF option. Lastly, additional comments indicated the importance of user buy-in with both the language used and customization of the system for the intended user base. This was perceived as very important for the ultimate adoption and diffusion of the system.

The participants also recorded their comments throughout the task completion portion of the usability testing process about dropped calls where the system would time out either because of too long a pause before the user response or a "processing error". Other comments focused on the meaning of some of the prompts in the system, such as the ordering of the menus or the unclear indication regarding when and how to respond to a particular prompt. One comment asked for more relevant information for the user population.

5.0 Discussion and Conclusions

The proxy evaluation was an extremely valuable approach to gain insights into the usability of the Spoken Web concept and the early version SW-CMP for healthcare management. Our analysis based on the survey responses, user comments, and

process diagrams generated from click-stream data indicated that proxy evaluators had mixed experiences with SW-CMP. They were generally happy with the consistency within the system and with the components of the system that were available to them. Consistent with prior research, they valued the Educational podcasts, although they suggested that the length of the menus associated with that section be modified for easier recognition. The ability of the users to access their health information, specifically having it spoken to them, elicited a positive response. The query section was also indicated as a valuable and well-functioning tool and resource for the users.

Much of the dissatisfaction was based on users' difficulties with unexpected closures while navigating through the system, which was reinforced both in the survey responses, comments, and process diagram analysis. Evaluators also had trouble accessing the system once they were forced out, which made the process more frustrating. This phenomenon may have been due to the logistical details of the call process, with SW-CMP based at IBM IRL in India while evaluators were located in the U.S. and using a call bridge to link the callers to the SW server free of charge. This bridge was used in both the development and testing stages; it was only problematic in the testing stages, possibly due to the larger number of users accessing the system simultaneously. This may have had an impact on the actual quality of the connection between the users and the system and users' opinions and perceptions. Further testing needs to be conducted under varying circumstances to test this problem in isolation.

The differences in satisfaction with some components of the system seemed to be partially based on the role of the evaluator. Evaluators with provider roles were generally more confident in the system and its potential to improve the management of the health conditions of patients; the patient evaluators thought that the system was not entirely user-friendly. It was recommended that the phone tree structure be simplified to facilitate memorization and intuitive navigation. These observations highlight the directions for improvement and are being implemented in the next iteration of SW-CMP. These improvements include more clearly-defined and articulated prompts for each feature; closing the loop by providing positive feedback to the user upon successful completion of a task; separating the group discussion forum from the secure patient-provider messaging feature; a ring-back feature to support medication adherence; and an analytics platform that uses the data streams

from the use of SW-CMP to analyze usability, performance, satisfaction and outcomes associated with the technology. While some of these improvements are very challenging to solve, such as voice recognition and language translation for better articulation, others such as medication adherence reminders, a discussion forum for social collaboration, and user feedback are currently being implemented. We hypothesize that these system improvements will allow a more interactive and meaningful process for end users, resulting in improved adoption and use.

Thus preliminary conclusions from this proxy evaluation about usability indicates that SW-CMP currently does not support an easy user experience, and this would be even more problematic for a more representative population; however, based on the comments and the survey results, it does promise opportunities for improved health and improved management of patients with chronic conditions. Thus, results were mixed, with users valuing the depth of information in SW-CMP for dealing with diabetes and its complications, ease of access once registered, intuitive nature after one or two tries, usefulness of podcast content, and customization potential for individual users. At the same time, lack of robustness, trouble dealing with phone trees, overly long Education section, and concerns with language recognition are challenges that are being addressed in ongoing studies, following which SW-CMP will be evaluated with actual end users.

Based on the feedback from current study, two ongoing studies with a re-engineered application, one with blind and visually impaired population and the second with low literacy and elderly population, are expected to provide unique and detailed insights into the specific requirements, new challenges and value of SW-CMP with the target demographics. In addition to the surveys, call log and usage data collection and voice files, these studies will employ think aloud protocols and observations and interviews of end users to elicit detailed information associated with using SW-CMP over an extended period of time.

SW-CMP was designed to meet the needs of a population that currently has limited or no access to any tool or resources to provide additional information about their health and wellness conditions outside the physician office, to support communication with providers, or to record personal health information. While SW-CMP is thus an improvement on the status quo for these individuals, we conclude that significant enhancements are needed on the usability and interaction schemes between the system and the individual to make it appropriate to meet the needs of the target users.

Ongoing and future research will explore these opportunities to better understand their requirements in order to create a meaningful experience for the patient while providing greater efficiency and cost effectiveness to the provider and address the health information and management needs of underserved and disadvantaged populations.

References

Agarwal SK, Kumar A, Nanavanti A and Rajput N. (2009), *Content creation and dissemination by-and-for users in rural areas*, ICTD '09 Proceedings of the 3rd International Conference on Information and Communication Technologies and Development, IEEE Press Piscataway, NJ, USA, pp. 56-65.

AHRQ. (2011), *National Healthcare Disparities Report,* Washington DC: US Dept. of Health and Human Services.

American Diabetes Association. (2012), *Diabetes Statistics.*

Baicker K, Cutler D and Song Z. (2010), *Workplace Wellness Programs Can Generate Savings*, Health Affairs, 29 (2): 304-311.

Bodenheimer T, Lorig K, Holman H and Grumbach K. (2003), *Patient self-management of chronic disease in primary care*, JAMA, 290(11):1455-1456.

Cohen R, Lemieux J, Schoenborn J and Mulligan T. (2012), *Medicare Advantage Chronic Special Needs Plan Boosted Primary Care, Reduced Hospital Use Among Diabetes Patients,* Health Affairs, 31, pp. 110-119.

Davis FD. (1989*), Perceived usefulness, perceived ease of use, and user acceptance of information technology*, MIS Quarterly, 13(3), pp.319-340.

Fisher EB, Boothroyd RI, Coufal MM, Baumann LC, Mbanya JC, Rotheram-Borus MJ, Sanguanprasit B and Tanasugarn C. (2012), *Peer Support For Self-Management Of Diabetes Improved Outcomes In International Settings,* Health Affairs, 31, pp. 130-139.

Goldman R, Sanchez-Hernandez M, Ross-Degnan D, Piette JD, Trinacty CM, Simon S. (2008), Developing an automated telephone diabetes intervention for health plan members. International Journal for Quality in Health Care, 20(4):264-70.

Gray B, Schuetz CA, Weng W, Peskin B, Rosner B and Lipner RS. (2012), *Physicians' Actions And Influence, Such As Aggressive Blood Pressure Control Greatly Improve The Health Of Diabetes Patients,* Health Affairs, 31, pp. 140-149.

Hall RF, Joseph DH and Scwartz-Barcott. (2003), *Overcoming obstacles to behavior change in diabetes self-management,* The Diabetes Educator, 29, pp. 303-311.

Hibbard JH and Greene J. (2013), *What the Evidence Shows about Patient Activation: Better Health Outcomes and Care Experiences; Fewer Data on Costs*, Health Affairs 32, no. 2 (2013): 207-14.

Hibbard JH, Greene J and Overton V. (2013), *Patients with Lower Activation Associated with Higher Costs; Delivery Systems Should Know Their Patients' Scores,* Health Affairs, 32, no. 2 (2013): 216-22.

Holman H and Lorig K. (2004), *Patient Self-Management: A Key to Effectiveness and Efficiency in Care of Chronic Disease*, Public Health Reports, 119, pp. 239-243.

International Diabetes Federation (IDF). (2011), *Global Burden*, Retrieved from <http://www.idf.org/diabetesatlas/5e/the-global-burden>.

ITU. (2012), http://www.itu.int/ITU-D/ict/statistics/material/pdf/2011, Statistical highlights June 2012.pdf.

Kumar A, Agarwal S and Manwani P. (2010), *The spoken web application framework- User generated content and service creation through low-end mobiles,* W4A2010- Technical.

Kumar A, Rajput N, Chakraborty D, Agarwal SK and Nanavati A. (2007), *WWTW: The World Wide Telecom Web*, In ACM SIGCOMM Workshop on NSDR.

Lerer A, Ward M and Amaransinghe S. (2010), *Evaluation of IVR Data Collection UIs for Untrained Rural Users*, ACM DEV'10.

Lewis JR. (1995), *IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use*, International Journal of Human-Computer Interaction, 7, pp. 57-78.

Mirza F, Norris T and Stockdate R. (2008), *Mobile technologies and the holistic management of chronic disease*, Health Informatics Journal, 14, pp. 309-321.

Nielsen J. (1994), *Heuristic evaluation*. In Nielsen, J., and Mack, R.L. (Eds.), Usability Inspection Methods, John Wiley & Sons, New York, NY.

Nielsen J. (1994), *Usability Engineering*, Morgan Kaufmann Publishers.

Norris P. (2000), *The worldwide digital divide: information poverty, the internet and development*, Annual Meeting of the Political Studies Association of the UK, London School of Economics and Political Science.

North Carolina Institute of Medicine (NCIOM). (2008), *Health Care Services for the Uninsured and Other Underserved Populations. A Technical Assistance Manual to Help Communities Create or Expand Health Care Safety Net Services*. Durham, NC.

Patnaik S, Brunskil E and Thies W. (2009), *Evaluating the accuracy of data collection on mobile phones: A study of forms, SMS, and voice.* In Proc. International Conference on Information and Communications Technologies and Development, pages 74-84.

Pickard AS and Knight SJ. (2005), *Proxy evaluation of health-related quality of life: a conceptual framework for understanding multiple proxy perspectives*, Med Care, 43, pp. 493-499.

QUIS™, Human Computer Interaction Lab, University of Maryland.

Schumacher RM and Lowry S. (2010), *NIST guide to the processes approach for improving the usability of electronic health records*, US Department of Commerce.

Sokol MC, McGuigan KA, Verbrugge RR and Epstein RS. (2005), *Impact of medication adherence on hospitalization rates and healthcare cost*, Medical Care, 43, pp. 521-530.

Suhm B. (2008), *IVR Usability Engineering Using Guidelines And Analyses Of End-to-End Calls*. Springer, US.

Sy D. (2007), *Adapting Usability Investigations for Agile User-Centered Design*, Journal of Usability Studies, 2(3), pp. 112-132.

Sommer A. (2009), *Getting What We Deserve: Health and Medical Care in America*, Baltimore: The Johns Hopkins University Press.

Tourangeau R, Couper MP and Steiger DM. (2003), *Humanizing self-administered surveys: experiments on social presence in web and IVR surveys*. Computers in Human Behavior, 19(1):1 - 24.

Wharton C and Lewis C. (1994), *The Role of Psychological Theory in Usability Inspection Methods*, In Usability Inspection Methods, J. Nielsen and R. Mack (Editors), John Wiley & Sons.

Wicks P, Massagli M, Frost J, Brownstein C, Okun S, Vaughan T, Bradley R, Heywood J. (2010), *Sharing Health Data for Better Outcomes on PatientsLikeMe*, J Med Internet Res, 12(2):e19.

World Health Organization: *Chronic disease*, accessed on January 10, 2013, http://www.who.int/topics/chronic_diseases/en.

World Health Organization (2011), *mHealth: New Horizons for Health Through Mobile Technologies*, Global Observatory for eHealth Series, Volume 3.

World Health Organization. (2012), *Diabetes Facts*.

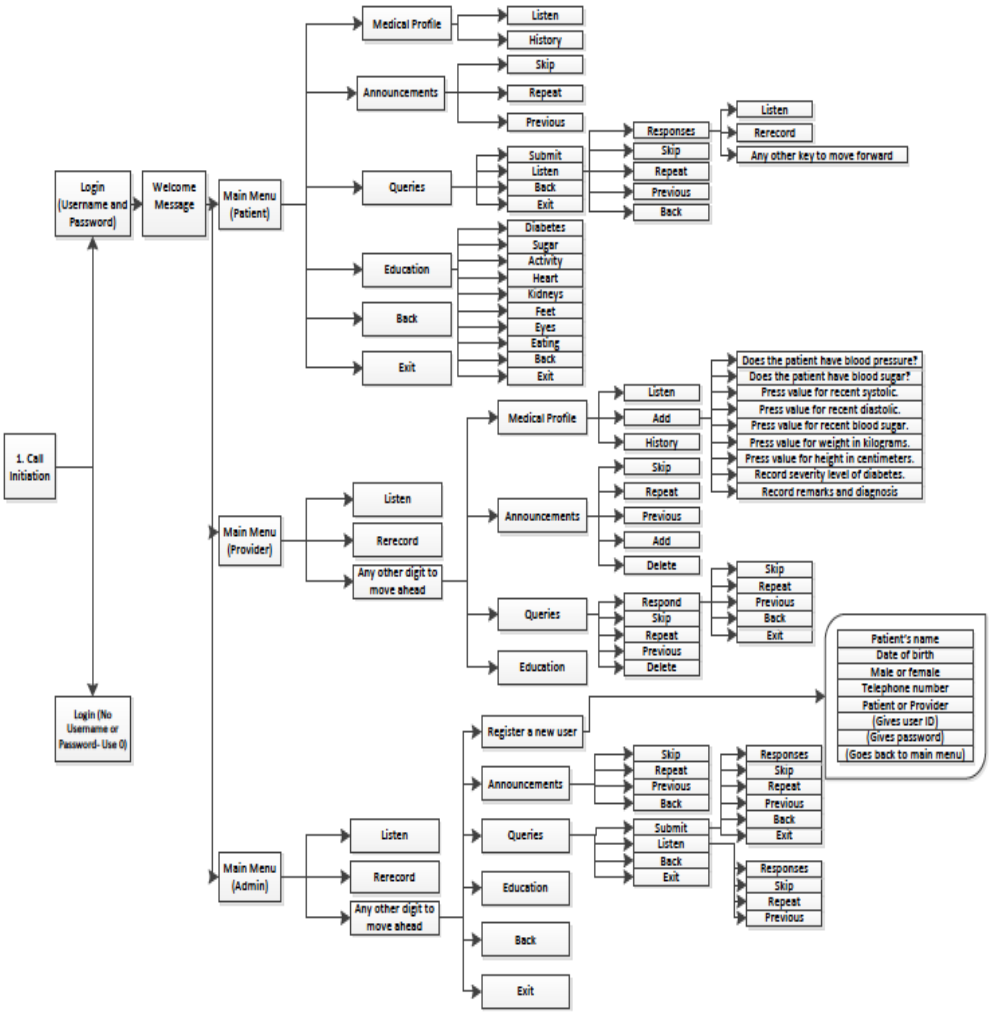

Figure 2. SW-CMP Navigation Chart

Question	Dissatisfied Response	Satisfied Response	All Evaluators			n	Providers Evaluators			n	Patients Evaluators			n
Responses (where D=dissatisfied, M=mediocre, S=satisfied)			D	M	S		D	M	S		D	M	S	
The terms of use throughout the SpokenWeb system was…	Inconsistent	Consistent	0.0%	0.0%	100.0%	8	0.0%	0.0%	100.0%	4	0.0%	0.0%	100.0%	4
The prompts for input were…	Confusing	Clear	12.5%	50.0%	37.5%	8	0.0%	50.0%	50.0%	4	25.0%	50.0%	25.0%	4
Performing tasks was straight forward…	Never	Always	28.6%	14.3%	57.1%	7	0.0%	0.0%	100.0%	3	50.0%	25.0%	25.0%	4
Correcting your mistakes was…	Difficult	Easy	42.9%	42.9%	14.3%	7	66.7%	0.0%	33.3%	3	25.0%	75.0%	0.0%	4
Was the ordering of menu options logical?	Bad	Good	0.0%	50.0%	37.5%	8	0.0%	0.0%	75.0%	4	0.0%	100.0%	0.0%	4
You would find it easy to get the system to do what you want it to do.	Unlikely	Likely	50.0%	50.0%	0.0%	8	50.0%	50.0%	0.0%	4	50.0%	50.0%	0.0%	4
I could use the SpokenWeb system successfully every time.	Strongly Disagree	Strongly Agree	50.0%	50.0%	0.0%	8	25.0%	75.0%	0.0%	4	75.0%	25.0%	0.0%	4
The SpokenWeb system appears to be designed for all levels of users.	Never	Always	14.3%	57.1%	28.6%	7	0.0%	66.7%	33.3%	3	25.0%	50.0%	25.0%	4
Learning to operate the SpokenWeb system was…	Difficult	Easy	25.0%	37.5%	37.5%	8	25.0%	25.0%	50.0%	4	25.0%	50.0%	25.0%	4
Remembering the names and use of commands was…	Difficult	Easy	37.5%	50.0%	12.5%	8	25.0%	50.0%	25.0%	4	50.0%	50.0%	0.0%	4
Would you say the SpokenWeb system is useful?	Strongly Disagree	Strongly Agree	12.5%	62.5%	25.0%	8	0.0%	50.0%	50.0%	4	25.0%	75.0%	0.0%	4
Would you say that the SpokenWeb system meets your needs?	Strongly Disagree	Strongly Agree	42.9%	42.9%	14.3%	7	33.3%	33.3%	33.3%	3	50.0%	50.0%	0.0%	4
Would you say the SpokenWeb system is user-friendly?	Strongly Disagree	Strongly Agree	25.0%	75.0%	0.0%	8	50.0%	50.0%	0.0%	4	0.0%	100.0%	0.0%	4
Overall, what did you think of the SpokenWeb system? Did you find it…	Terrible	Wonderful	12.5%	75.0%	12.5%	8	0.0%	75.0%	25.0%	4	25.0%	75.0%	0.0%	4
	Difficult	Easy	12.5%	87.5%	0.0%	8	25.0%	75.0%	0.0%	4	0.0%	100.0%	0.0%	4
	Frustrating	Satisfying	37.5%	62.5%	0.0%	8	25.0%	75.0%	0.0%	4	50.0%	50.0%	0.0%	4
I would recommend the SpokenWeb system to a friend.	Strongly Disagree	Strongly Agree	37.5%	25.0%	25.0%	8	0.0%	50.0%	50.0%	4	75.0%	0.0%	0.0%	4
I think that using the SpokenWeb system would improve my health.	Unlikely	Likely	37.5%	50.0%	12.5%	8	0.0%	75.0%	25.0%	4	75.0%	25.0%	0.0%	4
I think that using the SpokenWeb system would make it easier to keep track of my health information.	Unlikely	Likely	37.5%	37.5%	25.0%	8	0.0%	50.0%	50.0%	4	75.0%	25.0%	0.0%	4

Table 1. Usability Survey Results

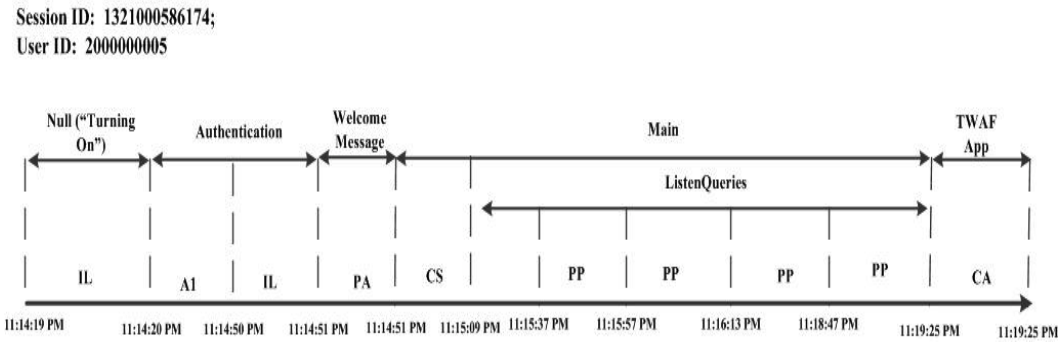

Figure 3. Process flow diagram of a call