Empowering ECAs in Handheld Devices

Ong Sing Goh, Arnold Depickere, Chun Che Fung, Kok Wai Wong
School of Information Technology
Murdoch University
Murdoch, Western Australia 6150
Email: {os.goh, a.depickere, l.fung, k.wong }@murdoch.edu.au

***Abstract* — The growth of mobile services on mobile phones or PDA has been significant in the recent years. This has opened up a new arena for many organizations to implement services on these mobile devices. In this paper, we propose an extension to our work in Artificial Life; of having our embodied conversational agents (ECAs) framework to be extended for mobile handheld devices. We present an overview and investigate new challenges on the implementation of the ECAs on the cross-platform architecture suitable for small end user devices. We demonstrated the proposed downsizing framework suitable for mobile devices with the application in the area of crisis communication.**

I. INTRODUCTION

Most research on embodied conversational agents (ECAs) has mainly focused on the conversations between humans and virtual agents on a web-based application. Interesting challenges and opportunities arise when the interlocutor is a physically embodied mobile agent— for example, a conversational robot or chatterbots. Our ECA can be seen as a digital spirit, capable of occupying various different bodies such as on the Internet, robot and handheld devices The portability of the agents and new miniaturized devices has make mobile services getting popular. Their ability to connect conveniently to networks anywhere anytime has also make mobile computing possible. During the past decade, one can observed that there are rapid advances in embodied conversational agents, spoken language technology, natural language processing, multimodal interfaces and mobile applications. All these have stimulated interest in a new class of *conversational interfaces* [1], [2], [3], [4] and [5]. Many researchers have also been observed in AI researches into natural language conversation [6], [7], [8], [9]. They have proposed different techniques and produced several natural language conversation systems. Every year they present their work by competing for the Turing Test [10].

This paper aims to address the issues of managing global crisis communication by introducing a crisis communication portal called Crisis Communication Network (CCNet). In particular, this paper focuses on two aspects of the system. First, we extend our ECA called Artificial Intelligence Neural-network Identity (AINI) framework suitable for mobile and handheld devices. This AINI delivers essential contents of news grabbed from trusted first sources online documents. Secondly, we implemented an application of the proposed ECA framework in mobile services for use in the area of crisis communication. The purpose of the AINI is to deliver essential information from trusted and updated sources and it is able to interact with its users through ECAs. The idea is to rely on a human-like communication approach thereby providing a sense of comfort and familiarity. Such "humanized" communication approach could be a favorable human computer interface on the web or mobile devices.

To achieve the above objective, we developed an intelligent agent software robot, AINI.. AINI has customized Artificial Intelligence Markup Language (AIML)[7] servable knowledge base being incorporated to serve as a real conversation software robot in the Crisis Communication Network (CCNet). Results from our simulation shows that the proposed framework could be a useful interface for mobile devices.

II. AINI ARCHITECTURE - SYSTEM OVERVIEW

This research project involves the establishment of a Crisis Communication Network (CCNet) portal. The objective is to use the AINI [11] embodied conversational agent as the basic architecture. The agent AINI can be seen as a digital spirit, capable of occupying and controlling a physical entity such as robot[12], or an embodied container, like an avatar in our conversational agent[13]. Our real-time prototype relies on distributed agent architecture designed specifically for the Web. A software agent, such as the conversation engine, multidomain knowledge model, multimodal human-computer communication interface and multilevel natural language query, communicates with one another via TCP/IP can be used. AINI is a conversation agent designed by the authors that is capable of having a meaningful conversation with users who interact with her. AINI is a software conversation

robot, which uses human-computer communication system. This is a combination of natural language processing and multimodal communication. A human user can communicate with the developed system using typed natural language conversation. The embodied conversation agent system will reply text-prompts or Text-to-Speech Synthesis together with appropriate facial-expressions.

For the purposes of this research, the application area chosen for designing the conversation agent is primarily grounded in an ability to communicate based upon scripting and/or artificial intelligence programming in the field of SARs epidemic crisis. A sample of the communication logs between a user and AINI in the CCNet portal is depicted in Fig. 6.

As shown in Fig. 1, AINI adopts a hybrid architecture that combines the utility of multidomain knowledge bases, multimodal interface and multilevel natural language query. Given a question, AINI first performs question analysis by extracting pertinent information to be used in query formulation, such as the Noun Phrases (NPs) and Verb Phrases (VPs). AINI employs an Internet three-tier, thin-client architecture that may be configured to work with any web application. It comprises of a data server layer, application layer and client layer. This Internet specific architecture offers a flexible solution to the unique implementation requirements of the AINI system.

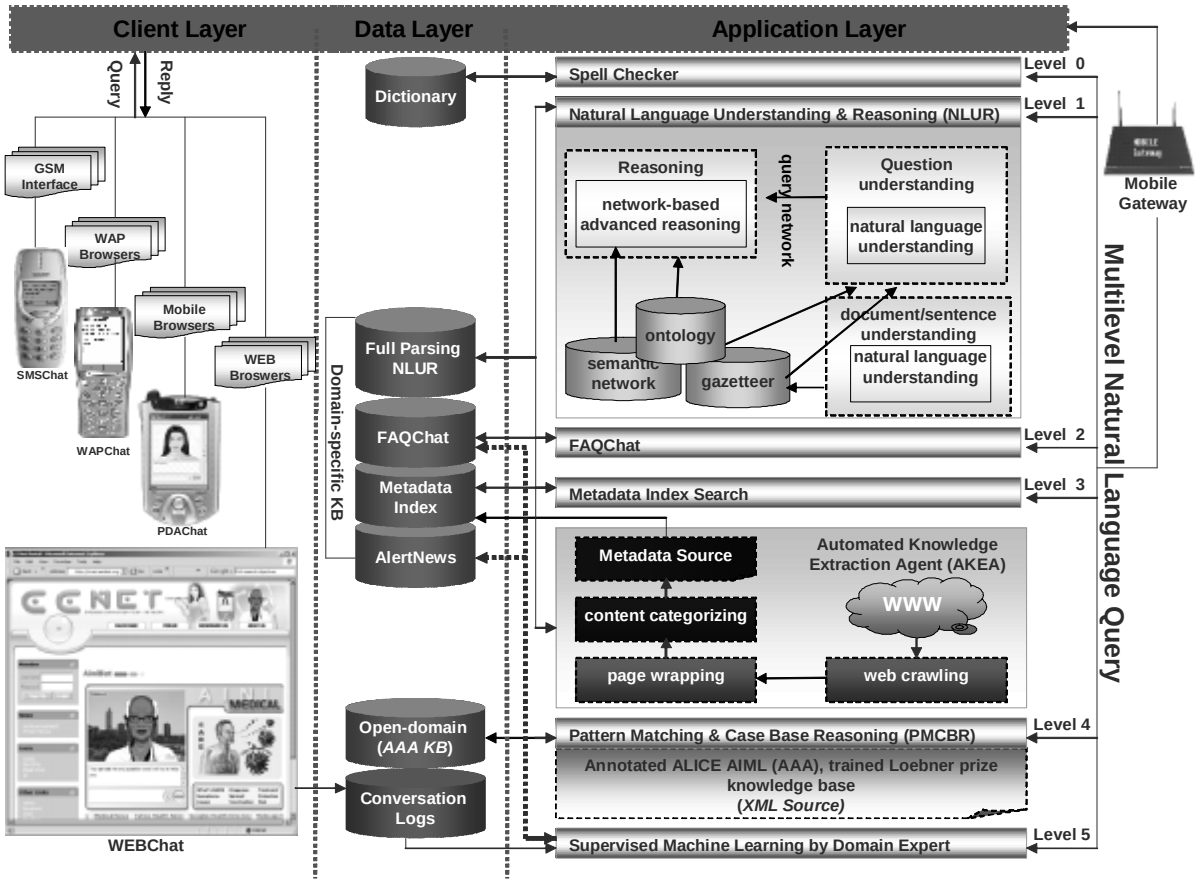

Fig. 1 AINI's Architecture in the CCNet Portal

A. *Data Layer*

The data server layer serves as storage for permanent data required by the system, where the epidemic knowledge bases are stored. These databases are Dictionary, Domain-Specific, Open Domain, AlertNews and conversation logs. The dictionary is an ***ispell*** which was first ran on TOPS-20 systems at MIT-AI lab[1]. Domain-Specific database is extracted by the Automated Knowledge Extraction Agent (AKEA)[14] which consists of Full Parsing Natural Language Understanding and Reasoning (NLUR), FAQChat and Metadata Index. AKEA was designed to establish the knowledge base for a global crisis communication system called CCNet. CCNet was proposed during the height of the SARs epidemic in 2003[15].

The Open-Domain database is taken from the existing award winner Turing Test [16]. This trained Knowledge Base is also called Annotated ALICE Artificial Intelligence Markup Language (AAA) [7, 16]

[1] http://www.mit.edu/afs/sipb/project/sipb-athena/src/ispell/

where the conversation logs reside. These web-enabled databases are accessible via the SQL query standard for database connectivity using MySQL database.

B. Application Layer

The AINI Server and Mobile Gateway are located in the application layer. WAP and SMS gateway[17] serve as mobile gateway used widely across the globe both for serving millions of short messages (SMS) and pushing WAP services. They function as the interconnection path between the client layer and data layer in the CCNet Portal. All communication with AINI is through a natural interface that uses a natural language understanding and reasoning; and text-to-speech via a 3D animated character called avatar. AINI's engine implements its decision making network based on the information it encounters in the knowledge bases. These decision-making capabilities make use of the XML specifications. The input and output of each module is an XML-encoded data structure that keeps track of the current computational state. These modules are conceptualized as transformations over this XML data structure. The system accepts questions and requests from users and processes the queries based on the information contained in AINI's knowledge base.

The application server layer handles the processing of logic and information requests. Here, one or more application servers are configured to compute the dialogue logic through the multilevel natural language query algorithm. In this layer we simulated goal-driven or top-down natural language query (NL-Query) approach as human's process their language [18]. The top-down approach seems to be a good model for explains how humans use their knowledge in conversation. After much literature search, we concluded that in the field of Natural Language Processing (NLP), the top-down approach is far the best approach. As shown in Fig. 1, our top-down NL-query approach consists of 6 level of queries, namely Spell Checker (Level 0), Full-discourse Natural Language Understanding and Reasoning (NLUR) (Level 1), FAQChat (Level 2), Metadata Index Search (Level 3), Pattern Matching and Case-based Reasoning (PMCBR) (Level 4) and Semi-Automated Matching Approach (Level 5).

C. Client Layer

The user interface resides in the thin-client layer and is used to support web-based and mobile services interface. For web-based interface, its employs Multimodal Agent Markup Language (MAML) interpreter or Microsoft SAPI to handle the users interface. MAML is a prototype multimodal markup language based on XML that enables animated presentation agents or avatars. It involves a talking virtual lifelike 3D agent character that is capable of involvement in a fairly meaningful conversation.

However for mobile devices interface with small screens, there is a limitation on the amount of information they can present at one time. Reading large amounts of information from such devices will require large amounts of scrolling. To reduce distraction, interactions, and potential information overload, a better way of presenting information might be through multilevel or hierarchical mechanisms[19]. Hence, chatting mode interface will be the best solution for mobile service. In addition, current wireless network service vendors have introduced a wide bandwidth telephone network, 3G communication [20], and it enhances the possibility of adapting a smartphone as a client in traditional distributed systems. On the PDA or SmartPhone, our system required Mobile Flash Player[21]. This Flash player is high performance, lightweight, highly expressive client runtime that delivers powerful and consistent user experiences across major operating systems, browsers, mobile phones and devices. For the WAP services, the application was embedded WAP browsers from vendors such as Openwave and Nokia. Our WAP mobile Internet solution is converted Web content to WML and adapted the presentation according to each individual client device using the User Agent Profile that defines characteristic and user preferences.

The conversation engine is implemented by open-source architecture employing Kannel Mobile gateway, PHP, Perl scripting language, Apache Server and knowledge base stored in a MySQL server.

III. CONVERSATION MODULE

AINI's conversation module which uses Artificial Intelligence and Natural Language Processing is an important underlying technology for the CCNet portal. By using a human-like software robot, users will have the impression of interacting with another human being who responds to their commands or queries. This is a more natural interface that human is familiar with. So, the main objective of AINI is to intelligently offer related information on various topics (e.g. SARs), where the service provided is in a virtual environment where no real live agents or specialists will be physically involved. This means AINI uses natural language parsing i.e. AIML and the AINI engine to search the predefined knowledge base as well as other data sources located in different systems via networking.

A. WebChat

The web chat sessions allow the users to interact in real-time with the software robot at the website. These web sessions can either be text-based or voice-based with a 3D animated character and Text-to-Speech

technology. Users are able to customize the interface, input questions and receive text responses directly from a website as shown in Fig. 2. Besides that, users can go through all the information on the website for the topics they are interested in. At the same time, they can place questions to receive more guidance on other topics.

Fig. 2 Webchat with personalization Interface for Human Interaction with AINI

A collaborative browser allows a portal to guide the users through the website of the organization by automatically "pushing" URLs and information from other websites to the user's browser. This not only facilitates communication between the software robot and users, but also allows the intelligence software robot to help users to locate specific information on their websites. This is because AINI is able to intelligently react to with the user's commands.

Users interact with AINI through the normal Internet ports, which are connected to AINI's knowledge base that provides WebGuide, WebTips and WebSearch engines [22]. The purpose of WebGuide is to guide users through the entire portal. It enables AINI to offer help without waiting for the surfer to ask. The WebTips engine, on the other hand, provides tips or hints to users. It is an intuitive feature that recommends links within the site. The WebSearch system is integrated to other search engines. It is a web tool, which can search for local sites as well as the Internet, online databases besides providing translations and other applications.

B. MobileChat

Mobile chatting module is implemented in a series of logical phases. We predict that text based agent-to-mobile chats with agent-to-Internet and Internet-to-mobile chats are going to be implemented. We add personalized characters with ECAs, and more game-like chatting environments to Mobile users.

Conversation chat applications were the first and most successful community applications for mobile services such as SMS, WAP/PML, and MMS extended by Web services. Hence, they can chat everywhere, at any time, with any device!

Fig. 3. SMSChat Interface

i. SMSChat and WAPChat

All the services are integrated as one mobile chat component for providing the latest alerts and information to users. Thus, the mobile chat is an alternative way where users can chat with AINI using SMS, MMS or GPRS services. The SMSChat services are on-subscribe mode, on-demand and on-alert news based services from the CCNet Portal. MMS is an SMS type service but with added image, voice, animation and many more features.

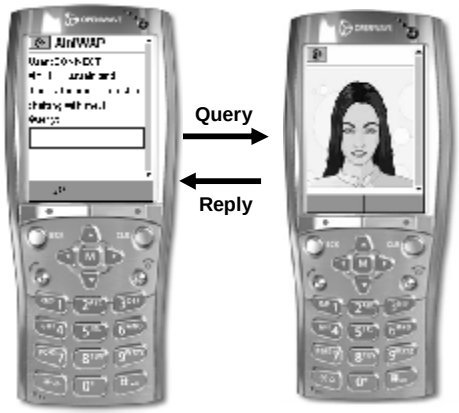

Fig. 4. WAPChat Interface

Therefore, depending on the display devices in use, the users will be able to view images of viruses, bacteria, infected cells and also X-Ray images. Meanwhile, the GPRS technology provides mobile web browsing functionality for accessing news and any other forms of data services on the CCNet Portal by connecting WAP gateway with given URL. WAPChat and SMSChat is the gateway to a new world of mobile data. It provides Text-based interactive information services and

applications from the screens of their mobile phones as shown in the Fig .3 and Fig. 4.

ii. PDAChat

The idea of developing AINI into Personal Digital Assistance (PDA)[2] is another approach of having a more human and personalized interface between a computer and human [23] [24]. The PDA chat with AINI

Fig. 5. PDAChat Interface

performs functions similar to that of by using webChat but mobile technology. It is a prototype designed to blend mobile technology with natural language to help humans interact more naturally with mobile devices. Here, the implementation of PDA chat with the knowledge base was designed using WiFi technology and powered by Microsoft Windows Mobile Technology [25], embedded with Pocket Internet Explorer as shown in the Fig. 5. While the computer-generated avatar of ECA is displayed on an PDA using Flash, the system uses a WiFi wireless connection to a server, where most of the work takes place.

IV. SMSCHAT – SELECTED CONVERSATION LOGS

In our experiment, AINI's domain knowledge model usually incorporates several knowledge domains, thus merging the expertise of one or more experts. Multiple domain knowledge merged into AINI's single domain knowledge would give the users the best conversation.

We predefined the Open-Domain and Domain-Specific in the data layer, while the Unknown-Domain is taken from a conversation log. Even though the conversation agent lets the users to direct the conversation in the Unknown-Domain, the agent will continue to remind and recall the users back to the current topic of the presentation. This is to convey and direct the users' attention back to its original Open-Domain or Domain-Specific state. However, the priority will be on the Domain-Specific. Therefore, the conversation agent will always give higher priority to Domain-Specific questions in an attempt to keep the user focused on the topic of the presentation.

To access these conversation and e-services through a mobile phone, users need to enter the given url into wap-enable browsers on the GPRS mobile phone that has internet access. On the mobile phone which unable to access WAP, the alternative is using SMS protocol. However, do note that the actual steps vary with different mobile make and model. An example of integration domain knowledge model and mobile services conversation between "Hussein" and "AINI" is shown in Figure 6.

```
1.User      : Hi
2.AINI      : Hello there. What is your name
3.User      : My name is Hussein.
4.AINI      : Nice to meet Hussein!
5.Hussein   : You still remember when was SARs
              first carried out into the
              world?
6.AINI      : Yes. SARs was first carried out
              into the world at large on 21
              February, at Metropole Hotel,
               Hong Kong.
7.Hussein   : When did Singapore mark on the
              SARs-free anniversary?
8.AINI      : Monday, 24 May 2004
9.Hussein   : Thanks for your information.
              When WHO declare world-free
              SARs?
10.AINI     :On July 5, 2003, the World
              Health Organization declared
              The world free of ongoing SARs
              Transmission.
11.Hussein  : Ok, now why not we talk about
               H5N1 Bird flu.
12.AINI     : I'm not sure how this related to
              SARs? Why don't you ask me
              more specific on SARs?
13.Hussein  : No, I am more interested on the
              current pandemic, H5N1?
14.AINI     : OK. I would do a search for it.
15.Hussein  : Thanks.
```

<Conversation continue and in the future, the Bird Flu H5N1 pandemic knowledge will be added by AKEA …>

Fig. 6. SMSchat Conversation Log

The conversation between *Hussein* and the *AINI* took place in the presentation on the topic of the SARs epidemic crisis using mobile services. From the = conversation logs, transition state in the line 1-4 used Open Domain and line 5-10 used Domain-specific. However in the line 11-13, user has control the conversation although the agent reminds the user of the topic of the current presentation. The transition state in

[2] PDAChat prototype available online at http://ainibot.murdoch.edu.au/ccnet/ainibot_pda.php

line 14-15 is under the Unknown-Domain state and the dialog set is taken from a pre-defined dialog set. The dialog contains *"I would do a search for it.", "Did I misunderstand your meaning?", "That's an interesting question. I'll come back to that in a minute.",* etc phrase will be monitored and submitted into the Unknown-Domain database. The author is freely to modify and enhanced the Unknown-Domain set to become a open domain or domain-specific knowledge base. This depends on the context or topic domain presentation.

V. DISCUSSION AND CONCLUSION

We have outlined the conceptual and practical basis for the development of the conversational agents for handheld devices. This will pave the way for more humanoid user interface based on human natural language technologies. Handheld devices provide an ideal platform for art and entertainment applications considering the growing number of mobile phone users world-wide. This will improve techniques for displaying content, interaction, conversation and the emergence of wireless and shared interaction among networked users. Such applications can captivate the users' attention through life-like features and the creation of personal attachment.

Based on this experiment, empowering ECAs on handheld devices shows interesting behavior in the natural conversation agent using mobile chatting. In this paper, we only worked on selected epidemic crisis websites when perform knowledge extraction for Domain-Specific knowledge on the server. Although we simulate the proxy conversation log that contains clients' requests, there is a possibility that new simulation result from other traces is different from the result referred to in this paper.

REFERENCES

[1] J. Rickel and W. L. Johnson, "Animated agents for procedural training in virtual reality: Perception, cognition and motor control.," *Applied Artificial Intelligence Journal*, vol. 13, pp. 343-382, 1998.

[2] J. Cassell, J. Sullivan, S. Prevost, and E. Churchill, *Embodied Conversational Agents*. Cambridge, MA: MIT Press, 2000.

[3] C. M. Karat, J. Vergo, and D. Nahamoo, "Conversational interface technologies," in *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications*, A. Sears, Ed. Mahwah, NJ: Lawrence Erlbaum Assoc., 2003, pp. 169-186.

[4] J. Lai and N. Yanlelovich, "Conversational speech interfaces," in *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications*, A. Sears, Ed. Mahwah, NJ: Lawrence Erlbaum Assoc, 2003, pp. 698-713.

[5] S. L. Oviatt, C. Darves, and R. Coulston, "Toward Adaptive Conversational Interfaces: Modelling Speech Convergence with Animated Personas," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 11, September 2004.

[6] J. Weizenbaum, "ELIZA - A computer program for the study of natural language communication between man and machine," *Communications of the ACM*, vol. 9, pp. 36-45, 1966.

[7] Alice, (2005),Artificial Linguistic Internet Computer Entity, [Online]. Available: http://www.alicebot.org

[8] B. Alfonsi, ""Sassy" Chatbot Wins with Wit," in *IEEE Intelligent Systems*, 2006, pp. 6-7.

[9] J. Hutchens and M. D. Alder, "Introducing MegaHAL," presented at The Human-Computer Communication Workshop, 1998.

[10] (2005),Loebner Prize, [Online]. Available: http://www.loebner.net/Prizef/loebner-prize.html

[11] O. S. Goh, C. C. Fung, and M. P. Lee, "Intelligent Agents for an Internet-based Global Crisis Communication System," *Journal of Technology Management and Entrepreneurship*, vol. 2, pp. 65-78, 2005.

[12] C. Uhrhan and O. S. Goh, "Features of a Mobile Personal Assistant Robot," presented at The International Conference on Robotics, Vision, Information and Signal Processing, IEEE-ROVISP 03, Penang, Malaysia, 2003.

[13] O. S. Goh, C. Ardil, W. Wong, and C. C. Fung, "A Black-box Approach for Response Quality Evaluation Conversational Agent System," *International Journal of Computational Intelligence*, vol. 3, pp. 195-203, 2006.

[14] O. S. Goh and C. C. Fung, "Automated Knowledge Extraction from Internet for a Crisis Communication Portal," in *First International Conference on Natural Computation*. Changsha, China: Lecture Notes in Computer Science (LNCS), 2005, pp. 1226-1235.

[15] O. S. Goh, C. C. Fung, A. Depickere, K. W. Wong, and W. Wilson, "Domain Knowledge Model for Embodied Conversation Agent," presented at the 3rd International Conference on Computational Intelligence, Robotics and Autonomous Systems (CIRAS 2005), Singapore, 2005.

[16] H. Loebner, (2003),Loebner Prize Gold Medal, [Online]. Available: http://www.loebner.net/Prizef/loebner-prize.html

[17] Kannel, (2005),Kannel: Open Source WAP and SMS gateway, [Online]. Available: http://www.kannel.org/

[18] O. S. Goh, A. Depickere, C. C. Fung, and K. W. Wong, "Top-down Natural Language Query Approach for Embodied Conversational Agent," presented at the International MultiConference of Engineers and Computer Scientists 2006, Hong Kong, 2006.

[19] S. Brewster, "Overcoming the Lack of Screen Spaces on Mobile Computers," vol. 6, pp. 188-205., 2002.

[20] GSMWorld, (2006),3GSM Statistics, [Online]. Available: http://www.gsmworld.com/technology/3g/statistics.shtml

[21] Adobe, (2006),Mobile and Devices, [Online]. Available: www.macromedia.com/mobile/

[22] O. S. Goh, S. Sahib, and R. Elangsegaran, "An Intelligent Virtual Financial Advisor System (IVFAS)," presented at 2nd International Conference on Neural Network and Computational Intelligence (NCI 2004), Zurich, Switzerland, 2004.

[23] HighBeam, (2003),Software robot for PDAs, [Online]. Available: http://www.highbeam.com/doc/1P1:84894246/Software+robot+for+PDAs.html

[24] H. Sulaiman, "Software robot for PDAs," in *Computimes, News Strait Times Press*. Malaysia, 2003.

[25] Microsoft, (2005),Microsoft Windows Mobile, [Online]. Available: http://www.microsoft.com/windowsmobile/