SmartWeb — Multimodal Interaction with Web Services

Norbert Reithinger, Simon Bergweiler, Anselm Blocher, Ralf Engel, Gerd Herzog, Alexander Pfalzgraf, Norbert Pfleger, Massimo Romanelli, Gerhard Sonnenberg, Daniel Sonntag, and Wolfgang Wahlster

DFKI GmbH, Stuhlsatzenhausweg 3, 66123 Saarbrücken, Germany,
firstname.lastname@dfki.de,
WWW home page: http://www.smartweb-project.de/

Abstract. SmartWeb aims to provide intuitive multimodal access to a rich selection of Web-based information services. The current SmartWeb prototypes are a smartphone client interface to the Semantic Web, an on-board car dialog system that gets update from the web, and a motorbike interface.

1 SmartWeb Handheld

The SmartWebconsortium started in 2004, and comprises fifteen partners from industry and academia in Germany. It aims at the development of a context-aware, mobile, and multimodal user interface to the Semantic Web [1]. In our main scenario, the user carries a smartphone as interaction device (see fig. 1 – other scenarios are the integration in a motorbike interaction device and in a car (see below). The user is able to pose multimodal open-domain questions in the context of a visit to a Football World Cup stadium in 2006. Using speech and gestures, she can ask for information about players, events, weather information and other services available through web services, Semantic Web pages, and the plain Web. The user can navigate through the result by speech and gestures on the device, and browses multimodally through alternative results.

The multimodal input is interpreted and transmitted using UMTS or wireless LAN to a back-end server system. The multimodal recognizers, the dialog system, and the Semantic Web access subsystems are located on this server. One server instance can serve multiple end-users. The response to a question is presented on the mobile device and rendered on its screen.

Key to our approach to multimodality in a question answering context are interoperating ontologies and a common data model for all knowledge-aware system modules. The ontological infrastructure of SmartWeb, the SWIntO (**S**mart**W**eb **Int**egrated **O**ntology), is based on a upper model ontology realised by merging well chosen concepts from two established foundational ontologies, DOLCE [4] and SUMO [5], into the SmartWeb foundational ontology SmartSUMO [6]. Domain specific knowledge (sportevent, navigation, web services) is defined in dedicated ontologies modeled as sub-ontologies of the SmartSUMO.

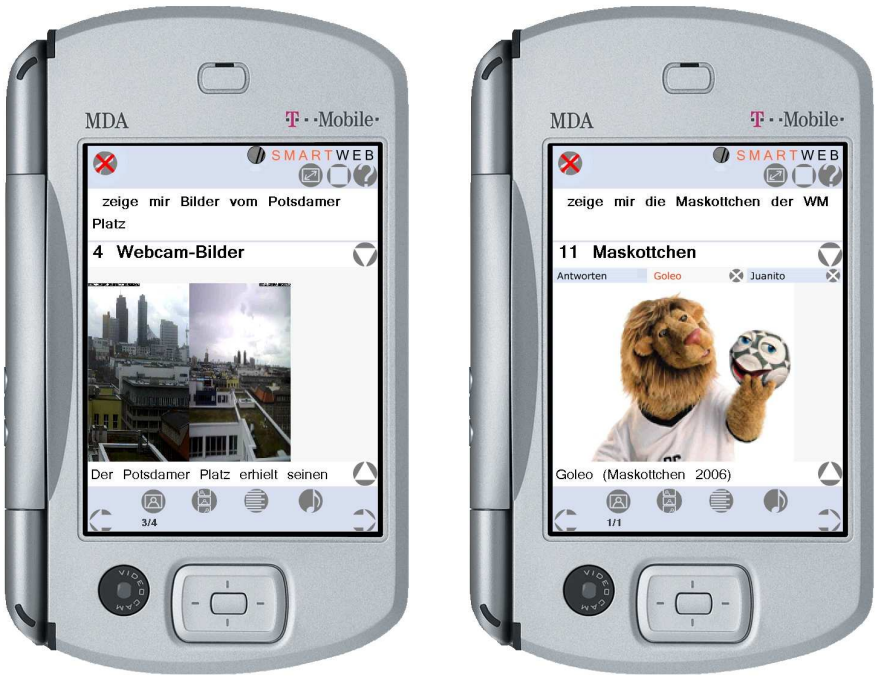

Fig. 1. SmartWeb handheld accessing a webcam service (left) and a knowledge server (right)

Additionally a discourse ontology (DiscOnto) provides question answering related concepts within SWIntOaswell as a media ontology (Smartmedia) provides representation of multimodal information. We make use of the W3C standard EMMA (Extensible MultiModal Annotation markup language)[1] for multimodal dialog management, that we extended for handling result feedback. The data exchange between components is RDF-based. [2]

Sympalog Voice Solutions, contributed their commercial speech dialog system platform that is being adapted and extended for multimodal interaction. The platform runs on Linux and contains all necessary functionalities for call handling over phone lines, sound-cards, or VoIP. It ships with client software for the PDA which runs Microsoft Windows Mobile 5. It takes care of the basic infrastructure, providing IP connectivity between server and client. The server side call-manager also provides stubs for speech recognition, synthesis and the dialog manager, and manages multi-user calls and base-level barge-in processing.

On top of this platform we developed the basic architecture that is shown in Figure 2. It consists of three basic processing blocks: the PDA client, the server system platform and dialog system. The client runs on the user's PDA.

[1] `http://www.w3.org/TR/emma/`
[2] http://www.w3.org/TR/rdf-primer/

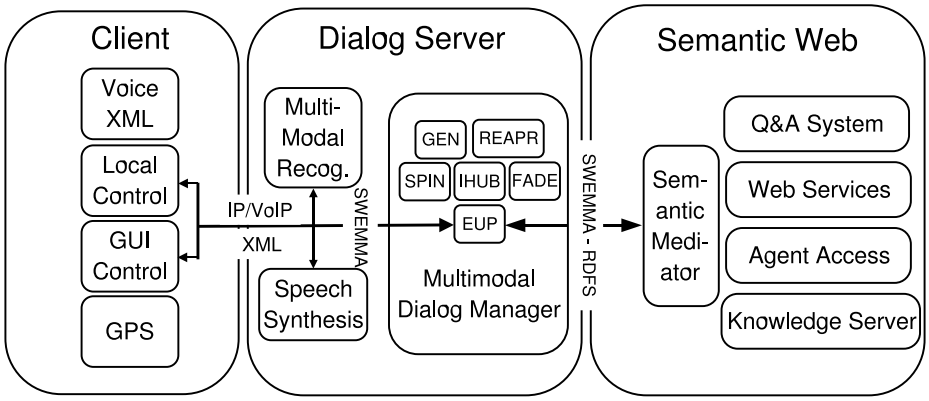

Fig. 2. The Architecture of SmartWeb Handheld.

It consist of a local Java based control unit which takes care of all I/O. It is connected to the GUI-controller which is realized using Macromedia Flash with ActionScript. Since the graphic realization on a small device is challenging, a professional environment like Flash provides for a straightforward and graphically satisfying development environment. The use of non-device specific programming languages and environments like Java and Actionscript/Flash adds another positive aspect: For development and test purposes we can run these components on a standard desktop computer. Additional modules are a local VoiceXML based dialog system for interaction during link downtimes, and camera and GPS based services connected to the controller.

On the server, the speech system platform instantiates one dialog server for each call, and connects the multimodal recognizer – currently speech recognition, later to be enhanced with head/gaze recognition –, speech synthesis and the dialog system.

The dialog system instances send the requests to the Semantic Services, which provide the umbrella for all different access methods to the Semantic Web the system supports. It consists of an open domain Q&A system, specialized web services, Semantic Web pages and a knowledge server. A component called Semantic Mediator encapsulates all these services.

2 SmartWeb – Car

SmartWeb car is a new mobile information system that interacts with drivers in natural language.

While you drive, the system searches the Internet for any potentially useful information. If you want to know which gas station in Dortmund has the lowest gas prices or how many goals Schalke 04 has scored, you can use SmartWeb Vehicle to retrieve this information from the system by means of voice input. This part of SmartWeb is being developed by the SmartWeb-consortium part-

ners from Siemens Corporate Technology Division (CT), DaimlerChrysler AG Research and Technology and from the Fraunhofer Institute for Device Architecture and Software Technology (FIRST).

The in-car system is divided into three basic areas: The first area includes the analysis of driver-relevant data from the Internet through a web crawler and its automatic processing into voice dialog systems. Since users are not supposed to address random queries to the on-board device, in their car is not a traditional search engine like those available on home PCs. The extraction is centered around services interesting to drivers. This process is supported by software that is capable of combining individual items of information into sentences and of breaking down spoken sentences into semantic units in order to understand them. The result is natural-language interactions between humans and computers where voice input no longer has to be restricted to a preset menu. The second area of the in-car system involves transmission of this information to the car via radio signals based on the new Digital Multimedia Broadcasting (DMB) digital transmission standard. The third step in the overall process is the actual interaction between driver and device in the car. Here care is taken to integrate the SmartWebcomponents seamlessly into already extsiting on-board information systems.

3 Acknowledgments

We would like to thank our partners in SmartWeb. This research was funded by the German Federal Ministry for Education and Research under grant number 01IMD01. The responsibility for this article lies with the authors.

References

1. Fensel, D., Hendler, J. A., Lieberman, H., and Wahlster, W.: Spinning the Semantic Web: Bringing the World Wide Web to Its Full Potential. MIT Press, 2003
2. Reithinger, N., Bergweiler, S.,Engel, R., Herzog, G., Pfleger, N., Romanelli, M., and Sonntag, D.: A Look Under the Hood - Design and Development of the First SmartWeb System Demonstrator. In: Proc. ICMI 2005, Trento, 2005 159 – 167
3. Wahlster, W.: Smartweb: Mobile applications of the semantic web. In: P. Dadam and M. Reichert, editors, *GI Jahrestagung 2004*, Springer, 2004 26–27
4. Gangemi, A., Guarino, N., Masolo, C., Oltramari, A., Schcneider, L.: Sweetening Ontologies with DOLCE. In Proc. of the 13th International Conference on Knowledge Engineering and Knowledge Management (EKAW02), volume 2473 of Lecture Notes in Computer Science, Sigünza, Spain, 2002, Oct. 1–4
5. Niles, I., Pease, A.: Towards a Standard Upper Ontology. In Proc. of the 2nd International Conference on Formal Ontology in Information Systems (FOIS-2001), C. Welty and B. Smith, Ogunquit, Maine, 2001, October 17–19
6. Cimiano, P., Eberhart, A., Hitzler, P., Oberle, D., Staab, S., Studer, S.: The smartweb foundational ontology. Technical report,Institute for Applied Informatics and Formal Description Methods (AIFB) University of Karlsruhe, SmartWeb Project, Karlsruhe, Germany, 2004