# TV AS A HUMAN INTERFACE FOR AMBIENT INTELLIGENCE ENVIRONMENTS

*Gorka Epelde, Xabier Valencia,*
Vicomtech-IK4, Spain
{gepelde, xvalencia} @vicomtech.org

*Julio Abascal*
University of the Basque Country, Spain
julio.abascal@ehu.es

*Unai Díaz*
Ingema, Spain
unai.diaz@ingema.es

*Ingo Zinnikus, Christian Husodo-Schulz*
DFKI, Germany
{Ingo.Zinnikus, chschulz}@dfki.de

## ABSTRACT

One of the challenges that Ambient Intelligent (AmI) faces is the provision of a usable interaction concept to its users, especially for those with less technical background. In this paper, we describe a new approach to integrate interactive services provided by an AmI environment with the television set, which is one of the most used interaction client in the home environment. An implementation of this approach has been carried out as a multimodal/multi-purpose natural human computer interface for elderly people, by creating adapted graphical user interfaces and navigation menus together with multimodal interaction (simplified TV remote control and voice interaction). In addition, this user interface can also be suited to other user groups. We have tested a prototype that adapts the videoconference and the information service with a group of 83 users. The results from the user tests show that the group found the prototype to be both satisfactory and efficient to use.

***Index Terms—*** Inclusive TV, Universal Remote Console, Multimodal interaction, Interactive TV, Interactive services

## 1. INTRODUCTION

As part of the Ambient Intelligence's vision, the environment has to seamlessly support the users in carrying out routine activities. The activities can range from home automation tasks to keeping up with friends in social networks. These tasks can be understood as services provided by an entity, integrated in the home environment itself and interacted through certain kind of User Interface (UI). For the redaction of this paper they will be generalized as interactive services.

Watching TV is one of the activities that take up most of people's leisure time [1], regardless of their technical skills and knowledge. Consequently, the TV has become a broadly extended electronic appliance, with a relevant place at home, and its remote control has become a common interaction device for most people. Consequently, the use of the TV as the most suitable UI to provide services adapted to each user in an AmI environment, seems very sensible. In this context, providing interactive services adapted to the user on the familiar device that is most frequently used by them in their spare time provides an interesting topic of research.

For many users, the integration of these services into a TV set provides the means being able to access online banking, e-health services, socializing via available social networks or making use of the services provided by their environment that otherwise would be inaccessible due to a lack of computer skills.

Up to now, the majority of research carried out on inclusive TV has attempted to integrate specific services, or to make the remote control of the TV accessible to groups of users with specific disabilities. As a result of this "craftwork" approach developed service adaptations are not reusable.

It is interesting that the largest efforts to develop natural interfaces for the TV set come from the interactive television area. In recent years, the iTV research community has progressed through the integration of interactive services within the TV. There is also an industrial commitment, to include such interactive functionalities in commercial products that have already been brought to market. However, most research and industrial developments have been targeted to the mainstream user, generally ignoring the accessibility barriers experienced by people with disabilities and elderly people. An approach to easily integrate interactive services with the TV, based on inclusive design, and one that would give universal access to the different user groups is still missing.

In this sense, the main contribution of this paper is a new approach to integrate all kind of interactive services (locally or remotely provided) with the TV set in a way that would allow personalizing the UI to the needs of each user group.

After this brief introduction, in Section 2 we have surveyed the current body of work related to the problem presented in the introduction. The next section introduces our approach to solve this problem. Afterwards, in section 4 we explain the evaluation method followed in the user tests.

Following, in section 5, we present the observed results. In section 6 we summarize the paper and draw our conclusions.

## 2. RELATED WORK

### 2.1. TV accessibility

The “Assessment of the Status of eAccessibility in Europe” report [2], analyzing both the accessibility of broadcast programs and end-user TV equipments, shows that TV accessibility is still far away from being implemented to its fullest extent. The main challenge of not excluding people from accessing any Digital TV’s services has been underlined by the iTV research community at [3].

From our point of view the three working areas directly related to making TV interaction accessible that must be addressed are: content accessibility, accessible remote control and accessible interactive services.

Regarding content accessibility, the main step was achieved by defining the accessibility services (audio description, close captioning and signing) for accessible broadcast content and the standardization of the technology used in the production, transport and rendering phases.

Subsequently, the accessible remote control topic covers the efforts of providing a means of remotely controlling basic TV functionalities such as channel up/down, volume up/down or turning the TV on/off, to all user groups.

Usability advances for different user groups have led TV remote controls to change their appearance.

The research community has also driven efforts to extend the capabilities of the remote control. These efforts are listed and classified Cesar et al. [4]. Some of the work that could help make the TV remote control more accessible is the use of objects such as pillows, gesture recognizers, speech interaction and dialogue systems [5], and the use of devices such as mobile phones and PDAs. These initiatives target specific scenarios and the developments cannot be easily integrated in future implementations.

In this sense, Epelde et al. [6] proposed an approach to make TV sets’ remote control accessible from design and to provide the ability to plug in different UIs developed for different controller technologies, such as those previously described.

With respect to the accessible interactive services research area, most of the work to date has been related to EPG applications. Some have extended the implemented Text to speech capabilities for the basic TV interaction activities to EPG applications. There is an initiative that integrates a paper based remote for interacting with the EPG application [7]. As in the TV remote control case, these efforts provide specific solutions to specific iTV applications in specific TV sets.

Besides these efforts, we have not found other initiatives that are focused on this area. In our opinion, it is necessary to define an approach that would allow the integration of services with different TV sets to meet different user needs.

## 3. ACCESSIBLE INTERACTIVE SERVICES

### 3.1. The URC framework

The Universal Remote Console (URC) framework [8] was published in 2008 as a 5-part international standard (ISO/IEC 24752). It defines a "user interface socket" (UI Socket) as the interaction point between a pluggable user interface and a target device or service. The framework includes "resource servers" as global market places for any kind of user interfaces and resources necessary for interacting with appliances, and services, to be shared amongst the user community.

Furthermore, the Universal Control Hub (UCH) is a gateway-oriented architecture for implementing the Universal Remote Console (URC) framework in the digital home [9].

### 3.2. Our approach

In this paper we present an approach that proposes the use of the URC framework in the form of a gateway-oriented architecture, UCH, to provide accessible interactive services to any TV set.

We adopted the URC Framework because it makes a clear separation between the service we want to access and the UI we want to use for accessing it. This way we are able to provide interactive services on any TV set. The TV sets that are used in this approach have the following requirements: they must implement a communication technology and a programmable user interface system. These TV sets can implement varying levels of accessibility features, depending on the user’s requirements, and they may have different form factors, ranging from a TV, to a Set-top box, or a PC based media center solution.

Regarding the services to be integrated, they can have different levels of openness: they may have proprietary access protocols, defined access APIs or web service specifications.

The services’ integration into the UCH architecture is achieved by means of defining the required XML files (UI Socket, Target Description, Target Resource Sheets) and implementing the corresponding code for the target adapter layer requirements (Target Discovery Module and Target Adapter) for each interactive service. For more information see [10].

Through the implementation of a UCH’s User Interface Protocol Manager (UIPM) we can implement any TV set’s compatible communication protocol. Using the UIPMs we have the ability to plug in the different pluggable UIs to TV sets.

After achieving the integration of the services with the UCH and the required UIPM, we are able to create UIs for any service. The approach also allows the creation of aggregated UIs composed of different services. At the same time, the UCH can be connected to different resource servers on the Internet that offer UIs and UCH integration modules that may be downloaded and used directly.

The figure 1 outlines our approach to provide accessible interactive services in TV sets. This figure shows different target services integrated using their own protocols and that are accessed from different TV sets. The resource server object reflects the option of using the UIs and integration modules downloaded directly from the Internet.

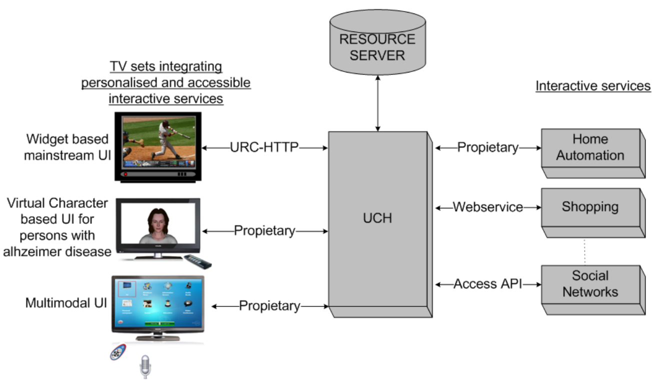


**Fig. 1.** Provision of accessible interactive services in TV sets.

Finally, the UCH implementation can be embedded into a consumer broadband router, into the TV itself, or with a more powerful UCH that has extended functionality, such as a dedicated PC.

## 4. IMPLEMENTATION

In the Vital project [11] we focused our approach's implementation on elderly users who are 60 years or older, with normal cognitive aging, and have an active life. With this idea in mind, we integrated those services that best fit the improvement to their quality of life, integrating them with mainstream society, and introducing them to new technologies through familiar interfaces.

The following services have been integrated to the UCH: videoconference as a social inclusion application, information service as a personalized information provider application, and audio book, educational content and p2p gaming (quizzes, chess) to enhance their leisure time and their cultural enjoyment.

With regard to the targeted TV set's UI, a multimodal interaction has been developed together with a simple and easy to navigate graphical user interface. The multimodal interaction includes both a simplified remote control and speech interaction modalities. The dialog system technology used in the implementation is explained in subsection 4.1.

The TV set's UI system is composed of a main menu with access to the different applications and the interfaces of the corresponding applications.

For the discussion of the evaluation and its results we will concentrate in the videoconference and the information service. In this way, these applications can be explained in more detail in the subsections 4.2 and 4.3.

### 4.1 Dialog System

We conceive the multimodal dialog system as a scalable and modular unit, which provides voice control over the applications integrated in the VITAL platform. The core component of the multimodal dialog system is constituted by the Ontology-based Dialog Platform framework (ODP) [12]. It provides an open architecture for building multimodal, task-oriented user interfaces that is in concordance with large parts in W3C's multimodal architecture proposal [13].

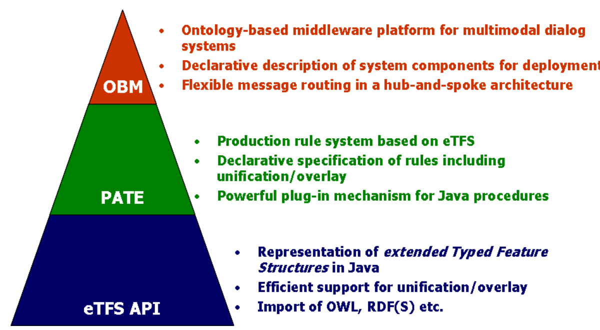


**Fig. 3.** The core building blocks of the ODP framework

The ODP framework itself is built up on blocks (see Figure 3) which stand for the basic dialogue system modules:

- The *Extended Typed Feature Structures* represents a data representation format which unifies the properties of RDF/RDFS (Resource Description Framework [14], and typed feature structures [15]). The encoding of the internal data as eTFS is accordant to the approach of ontology design in general, such that we use an **<object>** tag in order to denote a complex object of a certain type.
- *PATE* [16] provides a framework that supports the development of applications for multimodal dialogue systems. PATE's architecture is centered around the idea of three separated data storage facilities: (i) the *goalstack*, (ii) the *working memory*, and (iii) the *long-term memory*. The working memory is responsible for the activated instances so-called Working Memory Elements (WMEs), which are accessible for processing, i.e., rule applications. The long-term memory is responsible for the persistent storage for all instances of the type hierarchy the system has in the background. The purpose of the goal-stack is to represent the focus

of attention within the process of the system [17]. The placing of WMEs between the three data storage parts is organized by the activation value which changes in the processing flow and the effects taken by rule applications.

- *OBM* is a middleware platform and application programming framework for building multimodal dialogue systems. It is based on the eTFS API for system-wide data representation and PATE for implementing rule-based message routing. The OBM core functions as a server component that ties all modules and services together and maintains their interoperability. At run-time the server is responsible for managing the deployment of system modules.

As to enable voice interaction for the VITAL applications, the technical challenge was to first integrate the ODP framework into the UCH and second to provide the TV with the abstract presentation encoding the graphical content. In order to achieve a consistent context within the dialog system with respect to what is happening on the target side (e.g. Information Service Target) and on the User Interface side (TV) the internal state of the dialog system has to keep track and administrates the actions taken place at both ends. As we can see in figure 4, the ODP docks by means of the two different services at the UIPM Layer that speaks the URC-HTTP protocol, i.e., receiving and posting socket modifications from/to the targets and the protocol that can be interpreted by the TV client:

- **The Function Modeler** supplies the ODP with the information returned by the target that are converted into TFS objects and serve the Information State Module to update its own state.
- **The Presentation Planner** implements a service that invokes event handling on the UIPM. The TV client exclusively processes rendering information which is made ready by the UIPM. Also, in case the context information is not fully covered by the information retrieved from the target, VDS uses the rendering information as complementary input.

Besides the services, which is dealing with the communication to the TV and the backend services, different tasks within the multimodal dialog platform along the speech processing are allocated across multiple modules (see figure 4):

- **The Interpretation Manager** carries out natural language interpretation. For that purpose it processes the word lattice reflecting the user's vocal utterance with its semantic interpretation of the utterance. In particular, the natural language understanding component interprets the recognized spoken input of the user and converts it into instances of the ontology.

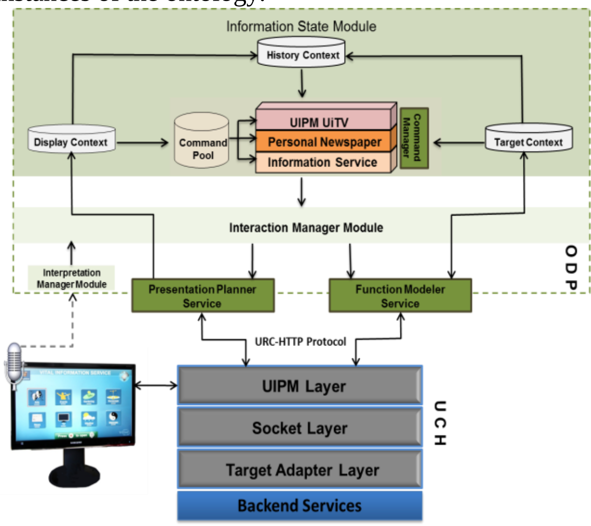


**Fig. 4** The data flow in the VDS architecture demonstrates how the Information State Module together with its building blocks is synchronized via two channels dedicated to the input coming from the User Interface (TV) and the backend services.

- **The Information State** stores and manages the ontology-based representation of all targets that represents the appropriate services on the backend side. Additionally, the Information State administrates and makes available a coherent representation of the displayed graphical content. The synchronization between the internal state of the dialog system and the backend components together with the content on the screen builds the basis for enabling access to multimodal interactive services. Depending on the state configuration, the Command Manager of the Information State Module retrieves the required command assigned to manipulate the states of the backend services or/and the GUI.
- **The Interaction Manager** has the task to propagate these commands and invoke the adapters that are able to speak the language of the URC-HTTP protocol, hence takes care of the information exchange with the target layer (backend services) and the presentation layer (TV). Typically when the user utters commands specific to the presentation layer, i.e., "Go to the Main Menu", first an ontological concept *SwitchApplication* is instantiated and then the interaction manager invokes the presentation planner to build and send the appropriate message to the UIPM.

Within the context of the tested prototype we can change between two input modalities and allow the user to alternate between voice interaction and the remote control. Possible utterances by the user are described by a grammar, maintained in a W3C standard compliant format. However, not only predefined speech input is accepted; the framework allows for loading new grammar entities on the fly. This is useful in the context of dynamic concept names (e.g. the title of a movie), which are created using information available from the web at runtime.

### 4.3 Information Service

An important concern we wanted to address was how to make the acquisition of information in the web easy to master. We believed that an application with such features integrated into the platform was of special interest for the elderly. To this end, we decided to avoid the use of web browsers, direct access to search engines (instead, the information service invokes a query on behalf of the user), and even hide the fact that the user accesses the internet at all. In order to accomplish transparency of the content for the elderly, we encapsulate the knowledge about the web site into an ontology tree [18].

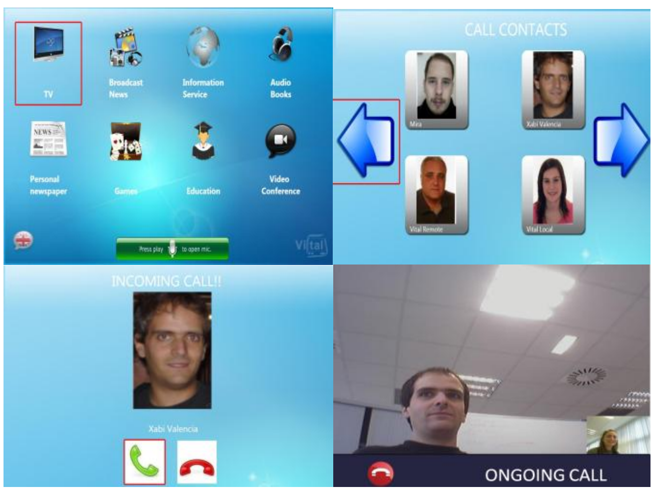


**Fig. 2.** Screenshots of the main menu and the three interfaces developed for the videoconference service.

Here, the ontology defines the conceptual relations in the domain. Furthermore, it assigns web pages to concepts and specifies the rules to extract the documents. In a second step, the ontology provides a description of content related to specific web pages. User preferences and interests of a specific user help to further restrict the space of concepts. Learning of users' interests is done by statistical evaluation of previous user behavior. Combining a probability approach and a vector space model, a personal recommendation service provides interesting documents which are instances of the favored concepts in the ontology tree. For instance two instantiated concepts that have been established by the users' preferences and can be seen on figure 3 are:

- **TV:** A personalized guide to the daily TV programme. The user can browse through all programmes split into categories (e.g. movies, sports, series and more).
- **Wellness:** Information about a healthy life style, suggestions for staying in good shape, news about advanced techniques in medicine.


**Fig 3.** A screenshot of the start page of the Vital Information Service is displayed. Preselected topic areas are distinguished by big icons and ordered by users' preference that may change after usage.

In the following sections, the tests of the two services that have been introduced are shown. Firstly, the evaluation methodology is introduced, and later, the results of the user tests are discussed.

## 5. EVALUATION METHOD

### 5.1. Participants

The sample recruited for the final evaluation of VITAL platform was composed of 83 participants, 19 male and 64 female, with an age ranging from 52 to 91 (x=73.68; sd=7.86) from the cities of Zarautz and San Sebastian, in the North of Spain. All users were attending elder associations in their respective cities. They had been living in their current location a mean of 46.88 years (sd=18.40), which identifies them as stable participants in their respective communities. Only 31.1% had no studies, while 45.9% had completed primary studies, 8.2% secondary studies, 11.5% had technical studies and a small 3.3% had completed university studies. They had been active workers throughout their lives, with a mean working life of 36.21 years (sd=13.03). 44.2% of the people from the sample were married, while 45.5% were widows.

### 5.2. Technical Set Up

Two laptops with a headset, a Weemote (simplified remote control), and their corresponding signal receptors were set in a room where the demonstration would take place on a group basis, connecting the main computer with the VITAL

Platform to a slide projector. One evaluator and one observer explained the whole procedure to the users in each testing site, gathered as a group in the room. They were administered a consent form, thus showing their acceptance to participate in the evaluation session. Afterwards, they were given the VITAL Questionnaire, which included the following sections: 1) sociodemographical data; 2) quality of family and social contacts; 3) leisure activities; 4) satisfaction with life, and 5) Specific evaluation of VITAL Platform services (here, questions about the system in general and individual applications in particular were asked).

### 5.3. Tests steps

After an approximate time of 30 minutes to fulfill sections 1 to 4 from the questionnaire, the main menu of VITAL was presented, and a brief explanation of its usage was given to the participants. Users were required to give written answers to questions related to the main menu, as well as to provide specific verbal feedback to what they were seeing on the screen and consulting the staff at any time. The same procedure was followed for each application (from videoconference to information service), thus showing the interface and the functioning (via Weemote) of each application. A demonstration of the interaction via voice (in English) was done. After all the services were presented, users were asked to discuss aloud any additional comments or feedback they would like to add. Then, the questionnaires were collected and users were thanked for their participation.

## 6. RESULTS

### 6.1 Overall impression about the system and main menu

The opinions about the system were very divided among those not having a clear statement, those thinking that it was a good application, and those reporting from the beginning that "this application may isolate people… it might isolate them in their homes". However, a majority of 68% thought it would be helpful in improving their social relationships, 72% thought it would help them to keep closer contact to their relatives, 75% thought it would help them to get closer with friends and 83.3% were confident in the idea that it would improve their quality of life.

Regarding the main menu interface, most of the sample (62.5%) found it pleasant or very pleasant, not being tiring for the eyes. All of them considered the interface was readable, with appropriate font size and color. The voice control demonstration worked well and participants were impressed. They expressed concerns regarding the use of headsets, the provision of the technology in other languages and the accuracy of the voice interaction with elderly users.

A table summarizing the results from the evaluation with users is included in table 1.

### 6.2 Videoconference Service

The layout of the videoconference was described as pleasant by 45.7% of the sample (45.7% said it was neutral, neither pleasant nor unpleasant). When they saw the way it worked on the demonstration session, 86.4% thought it was a useful application, and 43.4% would regularly use it (the others would rather continue with the regular phone). 62.5% would use it to talk to family, and 31.3% to both family and friends. It was a very well rated application ("it helps you keep in touch easier… it brings you closer to your relatives… in this way, I can see them").

Some of the participants were familiar with from of communication through Skype on a PC. However, it was considered an advantage that the videoconference system allowed being using the TV set for other purposes (watching a film, etc.) while it run in the background; the user could be just watching the TV and, in case of receiving an incoming call, the call would pop-up on the TV screen. This simplicity of use was highly appreciated by the users.

### 6.3 Information Service

62.5% described the layout as pleasant. After being exposed to its use, 100% considered it was useful, but only half of them would use it regularly. They liked the fact that local content was available but stressed that the content would have to be updated regularly in order for it to be useful. Many stated that they would continue with regular newspaper, and this kind of technologies may be good "for younger people". Some complained that the layout contrast was no good and that fonts should be made bigger, but they were told that this may be adjusted when using it on a TV set.

| |
|---|
| **Overall Impression about the system** |
| 68% found it helpful in improving their social relationships |
| 83.3% were confident it would improve their quality of life |
| 62.5% of the sample found the main menu interface pleasant |
| **Videoconference service** |
| 45.7% found the layout as pleasant (45.7% was neutral) |
| 86.4% thought was a useful application, 43.4% would regularly use it |
| **Information service** |
| 62.5% described the layout as pleasant |
| 100% considered it was useful, 50% would regularly use it |

**Table 1.** Summary of evaluation results

## 7. CONCLUSIONS

In this paper we have presented a new approach to integrate all kind of interactive services with the TV set in a way that allows personalizing the UI to the needs of each user group. The proposed approach is based on the ISO/IEC 24752 Universal Remote Console (URC) standard.

Our proposal for a "ordinary" TV user interface is based in the contributions provided by the interactive TV research. This approach allows access to interactive services from common TV sets, through the provision of personalized plug and play user interfaces that are rendered on the TV set. Following this approach permits an easy integration of new accessible services into our TV sets, including the services locally provided by intelligent environments.

At the same time, having the required modules available on a resource server on the Internet allows us to deploy and update our systems easily and opens a new market for service integrators and UIs developers.

An implementation of this approach has been carried out focused on the elderly. Services targeted on improving the elderly people's quality of life were integrated. With regard to the targeted TV set's pluggable UI, a multimodal interaction has been developed together with a simple and easy to navigate graphical user interface.

The user tests showed that the developed UI was well accepted and they thought that the developed concept could improve their social relationship and their quality of life.

## 8. REFERENCES


[1] Zillmann, D. and Vorderer, P.: Media Entertainment: The Psychology of Its Appeal (Lea's Communication Series). Lawrence Erlbaum Associates. (2000)

[2] European Commission. Assessment of the Status of eAccessibility in Europe. http://ec.europa.eu/information_ society/activities/einclusion/library/studies/meac_study

[3] Gill, J. and Perera, S.: Accessible Universal Design of Interactive Digital Television. In Proceedings of the 1st European Conference on Interactive Television: from Viewers to Actors?. EuroITV '01, 83-89. (2003)

[4] Cesar, P. and Chorianopoulos, K.: The Evolution of TV Systems, Content, and Users Toward Interactivity. Foundations and Trends® in Human–Computer Interaction 2 (4), 373-95. (2009)

[5] Berglund, A. and Johansson, P.: Using speech and dialogue for interactive TV navigation. Universal Access in the Information Society 3, (3-4), 224–238. (2004)

[6] Epelde, G., Carrasco, E., Zimmermann, G., Bund, J., Dubielzig, M., and Alexandersson, J.:URC based accessible TV. In Proceedings of the 7th European conference on European interactive television. EuroITV '09. ACM Press, New York, NY, 111–114. (2009)

[7] Berglund, A., Berglund, E., Larsson, A., and Bang, M.: Paper remote: an augmented television guide and remote control. Universal Access in the Information Society 4 (4), 300–327. (2006)

[8] ISO/IEC.: Information Technology - User Interfaces - Universal Remote Console (5 parts). ISO/IEC 24752. (2008)

[9] Zimmermann, G., and Vanderheiden, G.: The Universal Control Hub: An Open Platform for Remote User Interfaces in the Digital Home. In Proceedings of Human-Computer Interaction International 2007. (Beijing, P.R. China, July 22-27, 2007). HCI'07. Springer, Heidelberg, LNCS 4551, 1040-1049. (2007)

[10] URC Consortium. URC Technical Primer 1.0. http://myurc.org/TR/urc-tech-primer1.0/.

[11] Vital. Vital Assistance For The Elderly. http://www.ist-vital.org/.

[12] A. Pfalzgraf, N. Pfleger, J. Schehl, J. Steigner. ODP - Ontology-based Dialogue Platform. Technical Report, 2008, SemVox GmbH.

[13] McGuinness, D. L. and van Harmelen, F. OWL Web Ontology Language Overview. Technical report, W3C. http://www.w3.org/TR/owl-features/. 2004.

[14] Manola, F. and Miller, E. (2004). RDF Primer. Technical Report, W3C. http://www.w3c.org/Submission/OWL-S/.

[15] Carpenter B. (1992). *The Logic of Typed Feature Structures.* Cambridge University Press, Cambridge, England.

[16] Pfleger, N. and Schehl, J. (2006). Development of Advanced Dialog Systems with PATE. In *Proceedings of the International Conference on Spoken Language Processing (Interspeech 2006/ICSLP),* pages 1778–1781, Pittsburgh, PA.

[17] Anderson, J. R. and Lebiere, C. (1998). *The Atomic Components of Thought*. Erlbaum, Mahwah, NJ.

[18] Zinnikus, I., Fraser, C. & Fischer, K. (2004). An Ontology-Based Recommendation System. In Proceedings of Artificial Intelligence and Applications (AIA 2004), Innsbruck, Austria, 2004.