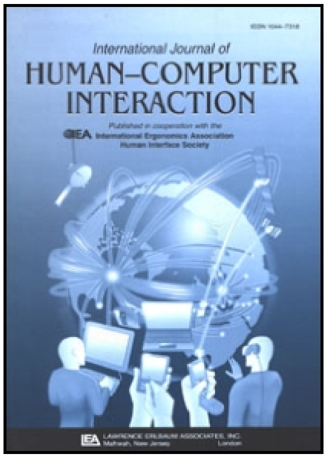

International Journal of
Human-Computer Interaction
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t775653655

Toward Advanced Social TV in a Cooperative Media Space

Tom Gross [a]; Mirko Fetter [a]; Thilo Paul-Stueve [a]
[a] Bauhaus-University, Weimar

Online Publication Date: 01 February 2008
To cite this Article: Gross, Tom, Fetter, Mirko and Paul-Stueve, Thilo (2008) 'Toward Advanced Social TV in a Cooperative Media Space', International Journal of Human-Computer Interaction, 24:2, 155 - 173
To link to this article: DOI: 10.1080/10447310701821491
URL: http://dx.doi.org/10.1080/10447310701821491

INTL. JOURNAL OF HUMAN–COMPUTER INTERACTION, *24*(2), 155–173, 2008

ISSN: 1044-7318 print / 1532-7590 online
DOI: 10.1080/10447310701821491

Taylor & Francis
Taylor & Francis Group

Toward Advanced Social TV in a Cooperative Media Space

Tom Gross, Mirko Fetter, and Thilo Paul-Stueve
Bauhaus-University, Weimar

Social TV systems provide groups of TV spectators with technical support for colocated and geographically distributed TV watching and social interaction. This article provides a systematic overview of the design space of Social TV and suggests a technical framework for flexible Social TV providing groups of TV spectators with sophisticated yet easy-to-use support for colocated and geographically distributed TV watching and social interaction. It sketches a scenario of advanced Social TV and then reviews previous concepts and systems as well as studies on Social TV to come to an in-depth presentation of design dimensions of Social TV. It introduces the Cooperative Media Space for Social TV as a technical platform for flexible support of advanced Social TV along the identified design dimensions.

1. INTRODUCTION

Social TV has a huge potential for making the group-viewing experience of TV spectators more enjoyable by providing novel means for advanced social interaction among spectators. The concept of Social TV can be defined as "the increasing integration of television and computer technology to support sociable, computer-mediated group viewing experiences" (Oehlberg, Ducheneaut, Thomton, Moore, & Nickell, 2006, p. 251).

In the following scenario the potential of Social TV is illustrated in four typical situations of advanced Social TV: virtual copresence of remote spectators while watching TV, communication with remote spectators while watching the same TV show, flexible adaptation of settings to groups of spectators while watching TV, and private conversations among group members.

Peter—a student from Germany currently on a student exchange in London—comes home from a busy day at university and wants to relax. He turns on his TV

We thank the members of the Cooperative Media Lab—especially Tareg Egla, and Christoph Oemig—for inspiring discussions on the concepts and guidance of the students for the implementation of PRIMI and Sens-ation. Special thanks also to Konstantinos Chorianopoulos and the anonymous reviewers for inspiring comments on earlier versions of this article.

Correspondence should be addressed to Tom Gross, Faculty of Media, Bauhaus-University Weimar, 99421, Weimar, Germany. E-mail: email@tomgross.net

set and activates the buddy list, which then appears on the right side of the TV screen and shows a list of all his friends who are online and watching TV at their homes (the top left of Figure 1 shows this situation of virtual co-presence of remote spectators).

Peter can see that a group of his friends are watching the news together in Sue's apartment in Berlin (cf. fourth row in the same buddy list). Peter asks them via a short text message whether they want to watch a movie together. They agree but first want to finish watching the news.

So Peter also tunes in to the news and is presented a screen with the TV news; on the right side of the screen a video window shows his friends, who are watching the same channel and sitting in Sue's apartment on a sofa, as well as another video window showing himself (the bottom left of Figure 1 shows this communication with remote spectators via audio and video). While watching the news, a discussion about some political topics arises in a conversation. As the discussion gets more vigorous, the platform adapts and lowers the volume of the news to allow a better conversation.

Meanwhile, Peter gets a notification from his buddy list, which is still running in the background, that his friend Paul wants to come over to visit him in his apartment in a few minutes. After Paul arrives, the friends in both places—in Peter's and in Sue's apartment—use an application called MovieMatch to select a movie everybody is interested in. By chance, Peter has the recommended movie in his personal DVD collection, and he starts broadcasting it to his friends in Berlin.

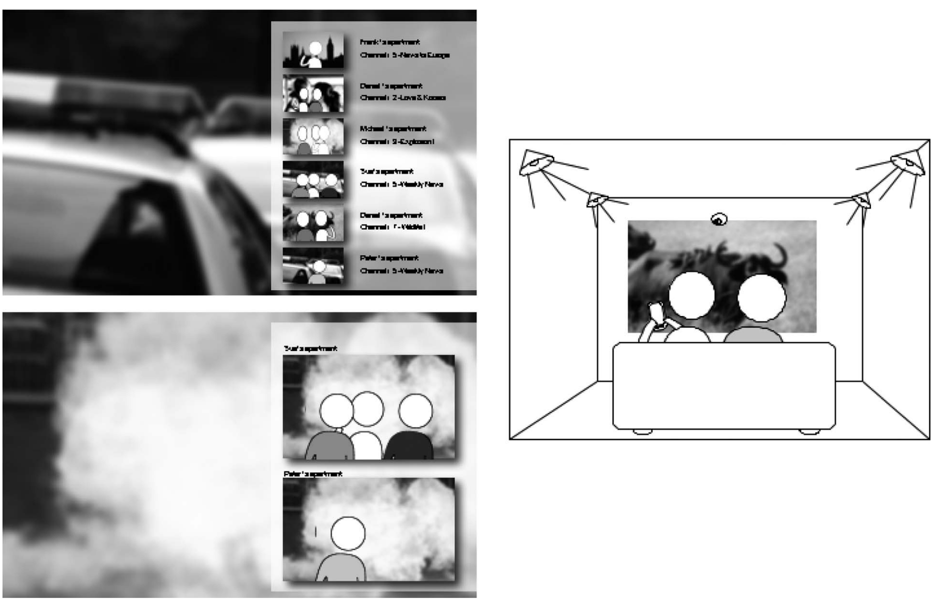

FIGURE 1 Mock-ups of the scenario: virtual co-presence of remote spectators while watching TV (top left); communication with remote spectators while watching the same TV show (bottom left); and flexible adaptation of settings to groups of spectators while watching TV at the same place (right).

To increase the immersion and the feeling of togetherness, the system now tries to match the lighting settings in the two rooms. It also adapts the TV settings like volume, luminance, and so forth, to best suit the personal preferences of all viewers (the right side of Figure 1 shows this flexible adaptation of settings to groups, including a device on top of the screen capturing the present persons and the current illumination of the room).

During the movie Paul starts chatting with one of his friends in Berlin about the main actress. Together they check an online movie database and talk about her other movies. They both use text messages on their personal digital assistant and mobile phone, respectively, because using the onscreen display and voice or video chat would probably disturb the other spectators (the right side of Figure 1 shows this private conversations among group members, including the personal digital assistant in Paul's hand).

In this article we want to depart from this scenario and elaborate on the design space for Social TV and report on a technical framework for advanced Social TV covering the dimensions identified. We depart from existing concepts and systems as well as studies of social action and interaction around home technology and TV watching and elaborate on the design space of Social TV and suggest a technical framework with sophisticated yet easy-to-use technical support for colocated and geographically distributed TV watching and social interaction. The following section glances at existing technical platforms for Social TV. Section 3 reports on existing studies and discusses design dimensions of Social TV. Section 4 introduces a technical framework for advanced Social TV and describes its two major components developed in our lab: the Platform for Research on Instant Messaging Infrastructures (PRIMI) for instant messaging and the Sens-ation platform for sensor-based environments. Finally, section 5 summarizes the contributions of this article and draws conclusions.

2. TECHNICAL CONCEPTS AND SYSTEMS FOR SOCIAL TV

Several articles have addressed issues of technological support for Social TV. They range from lightweight combinations of information and communication technology with TV technology to more sophisticated and integrated platforms. This section provides a brief overview of some influential work.

In several Human Connectedness systems Agamanolis (2006) explored the possibilities of new forms of distant social interaction triggered by mutual broadcasts among two sites. The Viper system offers producers of movies new types of narratives and easy editing of videos. The Reflexion video conferencing system provides a seamless integration of multiple video sources by extracting the image of the participants from the background and merging them together into one big life group image. It can also provide awareness on the speech activity level of the individual participants by putting active persons in the foreground and moving passive persons toward the background and continuously fading them out. The iCom system provides background awareness and chance encounters through permanent coupling of remote rooms. The tunA system is a mobile music-sharing tool. The Breakout for Two system is a prototype for sports over distance (Agamanolis, 2006). Finally, the Vision Television project (Bove & Agamanolis, 2006) uses similar foreground visualisations as suggested in our scenario just listed.

In a Social TV setup for a study Oehlberg et al. (2006) connected two distributed groups of TV spectators with an audio link, so all participants could hear the conversations in either room. The system is based on the Robust Audio Tool (Kirstein & O'Hanlon, 2006) running on a notebook below the TV sets.

Some integration of instant messaging and TV sets has been suggested. The Media Centre Buddies system integrates TV technology into an instant messaging application (Regan & Todd, 2004). The main purpose here was to allow multiple users to log into an instant messaging client and to experiment with notifications of incoming messages. Social Software for Set-Top Boxes is a thoughtful mock-up integrating instant messaging functionality into the TV (Coates, 2006). The author suggests that users can open a buddy list on the TV screen that shows all online buddies and the TV channels and shows they are currently watching. A pop-up alert window on the TV screen informs about buddies who switch channels. Furthermore, users can communicate with each other via text messages or audio and video conferences. Users can exchange information on schedules for upcoming recordings and exchange recordings if they miss recording themselves.

Finally, the Amigo TV system provides a technological platform for integrating content delivery, communities, and communication among spectators (Coppens, Vanparijs, & Handelkyn, 2005). The content of the broadcasts can be personalized by sharing personal photos and home videos. Community functionality allows users to meet online, see each other in buddy lists, and schedule meetings in online calendars. Communication support allows for voice, text, and video communication and expression through animated avatars. The Amigo TV system is based on IPTV (ITU, 2006) for personalized content, the eXtensible Messaging and Presence Protocol (XMPP; Day, Aggarwal, Mohr, & Vincent, 2000; Day, Rosenberg, & Sugano, 2000) for presence and buddy lists, and the Real-time Transport Protocol (Schulzrinne, Casner, Frederick, & Jacobson, 1993) for communication.

Overall the Amigo TV is probably among the most advanced Social TV systems. Yet it lacks a sensor-based approach for capturing the users' context and interaction with the environment and adapting accordingly.

3. THE DESIGN SPACE OF SOCIAL TV

In the last few years many concepts and systems for Social TV have been developed, and a considerable amount of studies mainly related to social issues of Social TV have been carried out. For a better overview of the challenges and opportunities in the design and development of Social TV systems this section elaborates on the design space of Social TV. To exemplify some opportunities it first introduces existing studies of Social TV. It then broadens up the scope to discuss general challenges and opportunities of by introducing design dimensions of Social TV.

3.1 Studies of Social Aspects of Home Technology, TV, and TV Watching

Several empirical studies in the private home with particular attention to the social interaction among people involved (typically family members and their visitors), and to the interaction of these people with their environment (typically the technical equipment and the space surrounding it), have been made and provide valuable hints with respect to the social aspects of home technology, TV, and TV watching.

O'Brien, Rodden, Rouncefield, and Hughes (1999) did a general study of the interaction among family members in private homes emerging around home technology that is very informative to designers and developers of interaction technology for private homes. The authors pointed out that "the contrast between work and domestic environments is not at all sharp" (p. 283). However, they also emphasized that the practices and activities in private homes can be very specific and technology not addressing these circumstances adequately can fail. The authors described some insightful ethnomethodologically informed ethnographic findings (Garfinkel, 1967). Here we can only provide two short examples that are relevant for our Social TV setting. First, they found a close coupling between the use of home technology and the space surrounding the technology. For instance, if a family has a TV set and a stereo set in their living room, the two typically mutually exclude each other. So the social arrangements regulate spatial and temporal structure and evolution of the occupation of space in the private home. Second, they found that each family member attributed specific sociability to the individual home technology. For instance, most family members would switch off the TV set when they receive visitors, whereas they would leave on the music of the radio.

Oehlberg et al. (2006) empirically studied the use of their Social TV system to get new insights with respect to social interaction among TV watchers. In their study they analyzed two settings: In the first setting the conversations in a colocated group of TV spectators was analyzed, and in a second setting the conversations in two remote groups of TV spectators connected with a permanent audio link were analyzed. It is interesting that the settings had low influence on the conversations—that is, the social interactions in the two settings had very similar nature and structure. However, the type of the TV show had a major impact on the conversations. For instance, the interaction during people-centered shows was more intense than during documentaries. Five different types of conversation contents could be identified: direct comments to the contents of the TV show, indirect comments with some relevance to the contents of the TV show, general social exchange about participants and their family and friends, organizational comments concerning the arrangement of the TV set and channel, and phatic responses about the participants' feelings. The authors came to the following concluding guidelines for the design of Social TV systems:

> support the proper timing of social interaction during group television viewing; minimise disruptions in the television program's flow; isolate exchanges that are beneficial to the group from side conversations and non-sequitur; allow viewers to move in and out of the audience smoothly; and avoid drawing viewers' attention away from the television screen. (Oehlberg et al., 2006, p. 256f)

3.2 Design Dimensions of Social TV

The aforementioned empirical studies describe typical interaction patterns of TV spectators using a Social TV system. They identify important aspects of typical functionality that Social TV systems should provide, of social settings, of spatial and temporal settings, and of adaptivity of the system to its users. In the following these aspects are discussed in detail.

Functionality. Social TV systems as reported in the aforementioned examples basically support the two core functions of group TV watching and of presence information and online communication, and optionally recommendations of channels, shows, or movies.

Group TV watching over multiple TV sets requires that the TV sets are connected and can exchange information on the channel or show currently running on each TV set. Most of the previously described systems support group TV watching (the Amigo TV system is certainly the most elaborated system for group TV watching; Coppens et al., 2005). Thereby, the individual group members in front of their TV sets can see either the same channel or show or different channels or shows. For multiuser applications in similar settings—that is, when multiple computers and the applications running on them are connected—several modes of coupling have been suggested (Dewan & Choudhary, 1991). The two most prominent are tight coupling and loose coupling; tight coupling refers to a system in which all users of the multiuser application work with the same data and have the same view on the data, and loose coupling refers to a system in which the users work with the same data but are free to move to different places within the data. Tight coupling according to the WYSIWIS (What-You-See-Is-What-I-See) principle typically entails the slaving of the view of all participants to one master—that is, whenever the master moves in the data, all slaves move automatically with it (Begole, Rosson, & Shaffer, 1999). The same principle can be applied to Social TV, where spectators might want to watch either different channels or the same channels or shows, or even might want to follow a specific spectator who shows them through the channels or shows. The latter mode mimics the situation of colocated group in which also one spectator typically has the remote control. In addition, systems can support the legal sharing of movies among users as exemplified in the aforementioned scenario.

Presence information and online communication are very important for the social exchange and coordination among remote spectators. Several of the previously described Social TV systems support presence information and online communication (especially the instant messaging integration of Regan & Todd, 2004). They have been widely explored in the computer-supported cooperative work (CSCW) and computer-mediated communication domain, where several concepts and systems for remote interaction have been developed within the last 2 decades. For instance, media spaces provide general awareness of the presence of remote parties and foster spontaneous and informal social interaction (e.g., Portholes [Dourish & Bly, 1992], RAVE [Gaver et al., 1992], Thunderwire [Ackermann, Hindus, Mainwaring, & Starr, 1997]). Thereby, technology should provide remote users with adequate awareness on the presence and availability of remote parties—in the context of Social TV especially with information on the availability for watching TV together or basic social interaction—without disturbing users too frequently from their primary activity (Gross & Prinz, 2004; Gross, Stary, & Totter, 2005). A central aspect to presence information is the users' free will—that is, only if users agree that their data should be captured and presented to other users (such as a user's presence in the system, the channels and shows currently being watched). Concerning online communication, basically the following media for communication can be used: text, audio, and video. Text communication can be easily stored in the context where it was produced and it allows parallel conversations, whereas audio and video communication is fast and easy but can conflict with the audio and video of the TV contents. Overall, the choice of communication media to support strongly depends on the context of use and base technology.

Besides group TV watching and presence information and online communication, Social TV systems can optionally support recommendations of shows and channels. Similar to recommender systems (Resnick & Varian, 1997) Social TV systems can automatically capture and store information on the shows and channels a user has seen and allow users to enter their personal ratings and comments. They can infer a profile of each user's preferences and present recommendations to individual users or shared recommendations to groups of users.

Social, spatial, and temporal settings. The social, spatial, and temporal settings mainly refer to the spectators per se as well as their geographical positions and the timing of their activities.

Social settings can include individual spectators, dyads of two spectators, groups of typically up to 15 spectators, large communities of much more than 15 spectators, or individuals or subgroups within groups of spectators. Furthermore, combinations are possible (such as an individual spectator in one place watching a show or channel with a group in a second place). Depending on the social setting, either concepts and mechanisms for small groups of typically up to 15 users can be found in CSCW and groupware literature (e.g., Borghoff & Schlichter, 2000; Marca & Bock, 1992), or in online and virtual community literature (e.g., Preece, 2000; Rheingold, 1996) for larger social settings. As Oehlberg et al. (2006) found in the study on Social TV use, the spectators involved in group TV watching have a considerable desire for interaction with each other—this should be considered when planning for large groups or communities.

Spatial settings can refer either to the absolute geographical position of the spectators or to the relative copresence of spectators (Jones, Grandhi, Terveen, & Whittaker, 2004). The absolute geographic position refers either to the static place where a spectator is or to the locations of a mobile spectator. The relative copresence of spectators gives insight on the distance between spectators: typically spectators can be either colocated at the same location or remote at different locations. Rodden and Blair (1991) identified two additional hybrid forms of colocation, where virtually colocated means that users are at different places and connected through low-fidelity information and communication technology (in a Social TV setting, this might be the case if one user is on the way to another user's apartment and uses the personal digital assistant to start negotiating the movie to be watched) and locally remote means that users have some high-bandwidth conferencing systems (in a Social TV setting this might be the case if two users have full-fledged audio and video connection between their apartments). As identified by O'Brien et al. (1999) in our study, home technology has specific implications on the use of the real space, and in particular the parallel use of different technologies in the same real space has its limitations.

Temporal settings can relate either to other spectators or to a specific show. Concerning other spectators the interactions of spectators can be either asynchronous, where the spectators use the Social TV system at different times, or synchronous, where the spectators use the Social TV system at the same time. Furthermore, systems can be of a mixed nature—according to Rodden and Blair (1991), mixed cooperative systems "contain elements of support for both synchronous and asynchronous cooperation" (p. 51). For Social TV systems, an example of mixed cooperation is when remote spectators can communicate with other online spectators via text chat in real-time and leave text messages to off-line spectators. Concerning a specific show, the actions of and interactions among spectators can happen before the show, such as negotiating the show to be watched together; during the show, such as commenting the current show; or after the show, such as rating and commenting the show just seen. As Oehlberg et al. (2006) found in their study the support of proper timing of the TV show and the social interaction during the TV shows is important. So flexible mechanisms for switching between different modes of coupling might be required.

Adaptivity. Usability in general is of central importance for the design and success of any systems (Hammond, Gross, & Wesson, 2002), and adaptivity in particular is a vital dimension in the design of Social TV systems. Adaptivity means that a system is "intended to behave according to the user's needs and preferences" (Sanches, Leggett, & Schnase, 1994, p. 369), whereas adaptability means that the system allows the users to actively change the system's behavior. As Oehlberg et al. (2006) discussed, the spectators typically had a considerable amount of organizational comments. Proper adaptivity might help reduce these types of conversations. Important design considerations for adaptivity are the procedure of processing the data, the users involved, and the application.

Systems supporting adaptivity typically use the following procedure: They capture and store various types of information about the user, such as shows and channels already or currently being watched; TV settings, such as volume and brightness; and environment settings, such as settings of the lights in the room where the user watches TV. They then typically use machine-learning algorithms to infer on users' preferences and needs (e.g., to detect a potential correlation between the sound level in the room and the preferred volume of the TV set). Finally they actuate—that is, adapt the application or environment according to their assumptions (e.g., adapt the volume of the TV set).

Adaptivity support can either address individual users or groups of users. For individual users the approach is straightforward: capturing and inferring on single-user data and changing the behavior. For groups of users a more complex approach is necessary typically involving the following steps: capturing the data of each user and establishing individual user profiles, inferring on the individual users' preferences and needs, searching for a common solution that fits all involved users' preferences and needs, and adapting the group settings accordingly.

Applications that are graphical user interface (GUI)-based typically capture the explicit user input from keyboard and mouse activities and adapt the behavior of the applications. Ubiquitous room-based environments capture the implicit input from user presence and activities in rooms, which is typically captured via sensors, and then adapt the room settings accordingly (Abowd, Mynatt, & Rodden, 2002). In the domain of ubiquitous computing, Weiser (1993) pointed out that computing technology will be increasing available and surrounding us and should therefore move to the background of our attention and be mostly invisible to its users. A core requirement to this end is calm technology—that is, the environment should capture the users' state and social context and actions in the environment and infer their needs and wishes rather than requiring users to explicitly give commands to the system (Abowd & Mynatt, 2000; Abowd et al., 2002). Clearly, when watching TV it is important that the interruptions of the spectators by the system are kept to a minimum.

The different design dimensions identified thus far are collected and presented in Table 1. Table 1 contains an overview of references. Overall there are references to all design dimensions—either concerning the whole dimension (with a reference entry in the first row) or concerning a specific point on a dimension (with a reference entry in one or more other cells). The lack of references in a cell can have two distinct reasons—either there is a lack of literature in this field (e.g., Social TV systems for mobile users) or there was simply not enough space in this article to go into the details of a well-established field (e.g., adaptivity in GUI applications).

4. SOCIAL TV SUPPORT IN A COOPERATIVE MEDIA SPACE

To develop technical concepts that aim to provide sophisticated support for Social TV yet stay in the background so the TV spectators can focus on the TV shows and the other spectators rather than on the technology per se, several technical requirements have to be met.

We suggest combining requirements and concepts from the CSCW domain with concepts from the ubiquitous computing domain. Ubiquitous environments (Weiser, 1991) do not burden the user with special requirements, such as particular steps of operation, but provide their functionality as features of the world. They respond to the needs and actions of the heterogeneous users by examining their social, spatial, and temporal context.

The Cooperative Media Space for Social TV (CoMeST) is a context-aware ubiquitous environment allowing natural interaction beyond traditional graphical user interfaces and supporting easy social interaction within colocated groups of users and among distributed groups. To support nonintrusive social interaction, information about the current activities of the users, their locations in space, their social environments, and their availability for conversations and for watching TV

Table 1: Overview of Design Dimensions of Social TV Systems, With References to Some Relevant Literature

Functionality	*Group TV Watching (cf. Coppens et al., 2005; Dewan & Choudhary, 1991)*	*Presence Info. / Communication (cf. Dourish & Bly, 1992; Regan & Todd 2004)*	*Recommendations (cf. Resnick & Varian, 1997)*		
Social settings: size (cf. Oehlberg et al., 2006)	Individuals	Dyads	Groups (cf. Borghoff & Schlichter, 2000; Marca & Bock, 1992)	Communities (cf. Preece, 2000; Rheingold, 1996)	Subgroups
Spatial settings: absolute (cf. Jones et al., 2004)	Static		Mobile		
Spatial settings: relative: (cf. Jones et al., 2004; Rodden & Blair, 1991)	Colocated (cf. O'Brien et al., 1999)	Remote	Virtually colocated	Locally remote	
Temporal settings: spectators (cf. Oehlberg et al., 2006; Rodden & Blair, 1991)	Asynchronous	Synchronous		Mixed	
Temporal settings: show spectators (cf. Oehlberg et al., 2006)	Before	During		After	
Adaptivity: procedure (cf. Oehlberg et al., 2006; Sanches et al., 1994)	Capturing	Inferring		Actuating	
Adaptivity: users	Individuals (cf. Abowd & Mynatt, 2000; Abowd et al., 2002; Weiser, 1993)		Groups		
Adaptivity: application	Graphical user interface application	Ubiquitous environment (cf. Abowd & Mynatt, 2000; Abowd et al., 2002; Weiser, 1993)			

is needed. Sensors capturing various data allow for inferring about users and users' context states. Through actuators the environment can support the users in their specific situation and balance between the needs of different users.

Subsequently we introduce CoMeST and its components for support of social interaction and for natural interfaces. First we introduce its components PRIMI and Sens-ation and then describe their integration into CoMeST.

4.1 The PRIMI Platform

PRIMI has been developed for the exploration of novel concepts in the domain of instant messaging with focus on easy support of social interaction in distributed groups (Gross & Oemig, 2005).

PRIMI facilitates the rapid development of the client-server infrastructures for instant messaging. New user interfaces can be easily integrated and custom and standard communication protocols are supported using a component-based approach founded on a plug-in architecture. It is generic and can integrate the functionality of other applications as well as be integrated into other applications. It may be distributed in a preconfigured manner as an end user application or configurable for expert users and developers.

PRIMI has a central kernel, the core of the platform, which is composed of two high-level and two low-level services. The high-level services are the ConnectorService and the AwarenessService. The ConnectorService is responsible for the integration of communication plug-in components. It is based on the low-level PluginService. The other high-level service, the AwarenessService, takes care of awareness information, which can be communicated with other applications. The low-level services are the PluginService and the LogService. The PluginService manages the loading and unloading of plug-in components. This can happen at start-up or at runtime via a network connection. The LogService allows for multi-level logging and wraps the logging of plug-in components; hence, plug-in component developers do not have to care about the logging details of PRIMI. There are two types of plug-in components in PRIMI. Communication plug-ins allow for the integration of different communication protocols, while user interface plug-ins allow for flexible user interface integration. An overview of PRIMI is given in Figure 2.

The PRIMI-AV extension of PRIMI supports text, audio, and audio–video chat. A topic-based communication mechanism allows users to join a topic-based conversation using different media. This can be several audio or video streams at once, or even still pictures. For instance, if a user wants to watch a movie with others, she can start a topic-based communication in which she broadcasts the title of the movie. Others can browse the current titles and join a topic-based communication, add their desired communication channels, and talk about the film with the others.

PRIMI-AV uses the Sens-ation platform (described next) to handle the communication management. Every PRIMI-AV client provides three software sensors. An availability status sensor feeds the active technical availability of different communication channels of a PRIMI user, such as streaming content or media

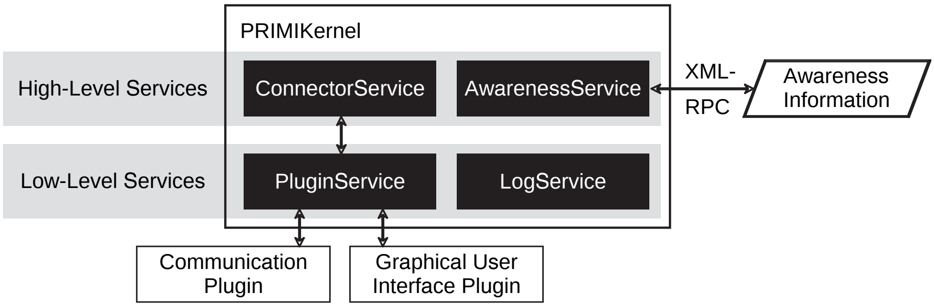

FIGURE 2 Schematic overview of the PRIMI software architecture.

data, to the Sens-ation platform. The communication request sensor reflects the user's intent to communicate with other users. It contains information about the addressees and the desired communication channels. The communication state sensor feeds the actual communication status of the PRIMI-AV client to the Sens-ation platform. To react on changes in the communication environment and on communication requests, every PRIMI-AV client subscribes to the software sensors of all other PRIMI-AV clients. PRIMI-AV streams audio and video content using the Real-time Transport Protocol (Schulzrinne et al., 2003) and uses the functionality of PRIMI for text-based communication via XMPP; exchange of raw data takes place via socket connections. An extended GUI provides access to the features of PRIMI-AV. An overview of the software architecture of the PRIMI-AV plug-in is given in Figure 3.

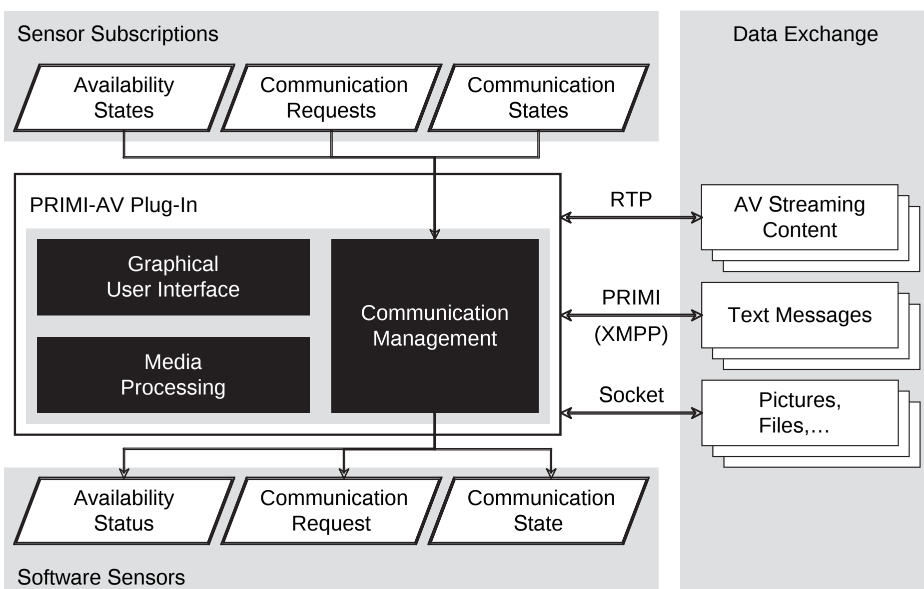

FIGURE 3 Schematic overview of the PRIMI-AV software architecture.

PRIMI is developed using the Java 2 Standard Edition 5.0 platform (J2SE 5.0) (Sun Microsystems Inc., 2006). Hence, it can be used on any operating system that provides a J2SE 5.0 runtime environment. The PRIMIKernel, a basic graphical user interface plugin, and a connector plugin are currently realized. The connector plug-in provides access to XMPP (Day, Aggarwal, et al., 2000; Day, Rosenberg, et al., 2000). It is based on the Smack API Library (Ignite Realtime, 2006). An XML-RPC (Winer, 1999) handler is integrated for the communication of awareness information. PRIMI-AV is optional and can be used to replace the basic graphical user interface plug-in of PRIMI. Like PRIMI it is developed using the Java 2 Standard Edition 5.0 platform (J2SE 5.0).

4.2 The Sens-ation Platform

Sens-ation is a service-oriented platform developed to free developers of context aware environments from technical constraints and details of sensors and actuators (Gross, Egla, & Marquardt, 2006). It can handle a variety of sensor and actuator hardware and software, at the same time abstracting from technical details of sensor data formats, protocols, and concurrency. Sensor values can be accessed from any location and various applications in nearly real time while a history of all earlier sensor values is provided. As a generic and service-oriented platform, Sens-ation can be used as middleware between applications and the sensor and actuator hardware or act on its own by directly controlling actuators dependent on current and past sensor values. Figure 4 shows an overview of the Sens-ation platform.

The communication with the sensors is realized through adapters. These are responsible for the introduction of new sensors to the platform and for the preprocessing of the sensor data. They abstract the technical details by translating raw sensor data into EventXML format, an XML-based format for the description of various types of sensor events. The adapters buffer sensor events, notify the platform, and send EventXML data to Sens-ation.

Handlers in Sens-ation take care of the different sensors. They register new sensor locations and sensor types to the platform. Sensor ports route incoming sensor data to the sensor data storage, the inference engines, subscription managers, and the gateways. Historic sensor data are stored in a database. Every incoming sensor event is stored to enable inferring using past and present sensor data.

At the moment, two mechanisms process sensor data: the inference engine and the subscription manager. Custom inference on data may be integrated in the Sens-ation platform as inference engine. Via subscriptions defined on application level, it is possible to filter sensor event data. These data are directly forwarded to the subscribing applications via the publishing mechanism of Sens-ation.

The sensor event data can be accessed via several different gateways. Sens-ation provides Web services, basic XML-RPC, socket connections via TCP/IP, a CGI interface, and an HTML interface.

The current Sens-ation version 3.1 is realized with J2SE 5.0, Apache Axis (The Apache Software Foundation, 2005), a MySQL database (MySQL AB, 2006), and

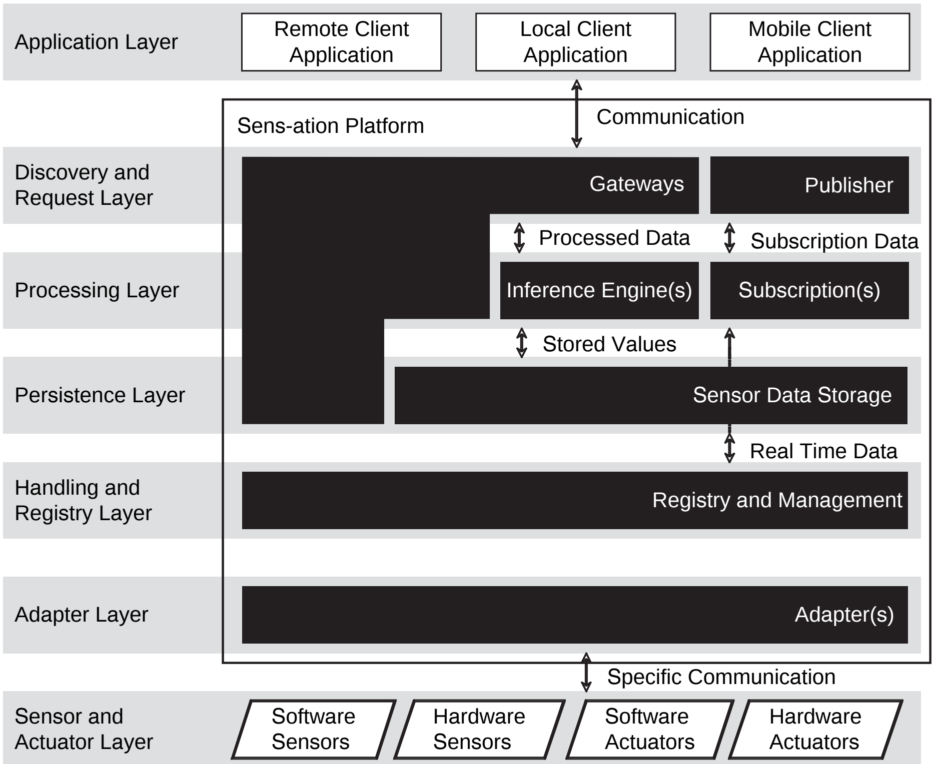

FIGURE 4 Schematic overview of the Sens-ation software architecture.

standard protocols. It is therefore highly portable and can be easily integrated with other platforms.

4.3 The Integrated CoMeST Platform

All technologies—the PRIMI platform and its PRIMI-AV extension as well as the Sens-ation platform—have been integrated into CoMeST. Leveraging the described platforms, components and technologies we are able to build up a technological framework converging in CoMeST for advanced Social TV according to the aforementioned scenario. Figure 5 shows the overall CoMeST software architecture.

As the PRIMI platform offers a solid base for context-aware social software, a huge part of the functionality can be integrated into PRIMI in form of a SocialTV GUI plug-in. The SocialTV GUI plug-in is tightly connected with PRIMI via the PluginService. It provides the basic TV functionality. For this purpose the Receiver Module allows the input of multiple different video signals and formats. Such a module can be arbitrarily expanded from using analog or digital TV via TV cards over streaming media from Internet video platforms (e.g., YouTube

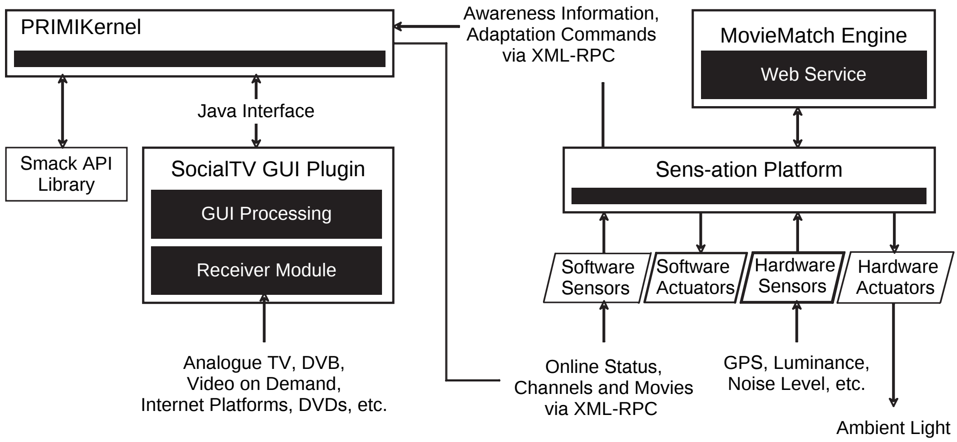

FIGURE 5 Schematic overview of the CoMeST software architecture.

[2006]), video on-demand services (e.g., NTL [NTL Group Ltd., 2006]), or personal online video recorders up to the personal video collection or WebCams as a source. The GUI Processor takes the selected video signal, presents it on the screen, and shows presence and context information of buddies enabling text chats or showing live video streams of the other viewers overlaid on the right side of the screen. In addition, the GUI Processor acts as a controller delegating the user inputs to the modules responsible for the TV program input as well as the parts ensuring communication and awareness.

The MovieMatch Engine is an extension of the Sens-ation platform, which can create substantial user profiles from the persistent data. These profiles can be used to serve as a data source for a recommender system. Based on the history of watched movies from all users involved in a certain social gathering, the software can suggest a movie that most likely fits the group's needs. MovieMatch is rooted in the MatchBase concept and technology for analyzing user behavior, inferring user profiles, and comparing user profiles (Gross, Braun, & Krause, 2006).

The communication between the individual components of this technological framework works as follows. The PRIMIKernel communicates with the SocialTV GUI plug-in via the PRIMI plug-in mechanism based on Java interfaces. The PRIMIKernel receives awareness information and adaptation commands from the Sens-ation Platform via XML-RPC calls; in return the PRIMIKernel acts as a software sensor and sends information that it receives from the users in the GUI, such as online states, channels, and movies watched, back to Sens-ation via XML-RPC. And the Sens-ation Platform communicates with the MovieMatch Engine via a Web Service gateway. Finally, the communication between the individual installations of the client software at the different sites (i.e., apartments of spectators) is based on the Smack API Library implementation of XMPP.

The data on all users including their actions and preferences are stored in a central instance of the Sensor Data Storage. These data include past and present information on users' login and logout and thus presence in the system, history of channels and shows already seen, ratings and comments to channels and shows, settings of the TV set such as volume and luminance, settings to rooms such as lightning, and so forth. All these data are available and used for only inference and adaptation within the CoMeST platform and cannot be accessed from outside. So, for instance, the history of shows and channels a user has seen are used only to infer each user's interest profile, and the details of the shows and channels seen and rated cannot be directly accessed by other users or systems.

5. DISCUSSION AND CONCLUSIONS

In this article we presented a brief survey of existing work on Social TV. We have derived design dimensions for Social TV support. CoMeST is intended to provide flexible support for Social TV both for the particular scenario previously depicted and along the general design dimensions elaborated earlier.

With respect to the previous scenario, the PRIMIKernel is the place where all information runs together. So when Peter checks the presence status of his buddies, or is chatting with his friends on the display or via the mobile client, it is always the ConnectorService in charge. Via the AwarenessService it is possible to get concrete information about the context of his friends. The AwarenessService receives the information and the commands as requests for adapting the GUI for the users from the Sens-ation platform.

The Sens-ation platform incorporates various hardware and software sensors for capturing information about users. For instance, in extension of the previous scenario, Paul's position can be tracked with a GPS hardware sensor, which communicates the position data via the Adapters to the Sens-ation platform. So it is possible for Peter to see Paul's position. Peter has just to subscribe for a notification event to be noticed when Paul is in a certain range of his house (if Paul has granted access to these data). Using software sensors, it is also possible to detect the title of the video, TV channel, or movie currently watched in the TV application. In a similar way luminance or noise data collected by hardware video and audio sensors can trigger hardware and software actuators to adjust the light situation or control the volume as described in the scenario. By using the sensor data history collected over multiple, connected Sens-ation instances and data-mining techniques, it is possible to adapt the environment to best fit all users' preferences.

From a broader perspective and along the design dimensions previously described, CoMeST is intended to provide flexible support for Social TV as follows:

- It captures the position and interaction of the users with the environment and with each other and *adapts its settings and behavior* to the evolving needs of the local parties. Besides the functionality already described, it can capture who is present and log in arriving users and log out leaving users automatically. Through a multiuser mode in the instant messaging functionality, it can manage any number of synchronously logged-in users on one machine.
- It *provides unobtrusive distant interaction* through subtle cues on the presence and online states of remote TV buddies as well as on incoming text messages

or audio and video chat requests (e.g., through subtle ambient sounds, or through subtle changes of the lights).
- It *protects the privacy of the TV spectators* involved by allowing users to specify their preferences if they want to get logged in and out automatically and about what to share with others (e.g., the degree of personal details on presence; the degree of the details of the channel and show currently being watched; the degree of the details of the persistent storage of data on persons, channels, shows, etc.).

We are currently working on the completion of the implementation of the CoMeST technological framework. In the current state the PRIMI platform (Gross & Oemig, 2005) including privacy protection through selective information disclosure (Gross & Oemig, 2006), and the Sens-ation platform (Gross, Egla, et al., 2006) including all necessary sensors and actuators (Gross & Marquardt, in press), and MatchBase including user profiling and comparison (Gross, Braun, et al., 2006), have all been implemented as previously described. We have developed concepts for the SocialTV GUI plug-in, and the MovieMatch Engine, but their implementation and integration is not yet completed.

Overall it is clear that the scenario in mind has an influence on the design dimensions considered and the concept of the technical platform designed and developed. The particular focus of this article was to contribute to the analysis and design of viewer- and context-sensitive group TV watching in both colocated and distributed settings. Other issues such as technical challenges and solutions for broadcasting, digital rights management, and so forth, are certainly also of vital importance for Social TV systems but had to be left out because of a lack of space.

REFERENCES

Abowd, G. D., & Mynatt, E. (2000). Charting past, present, and future research in ubiquitous computing. *ACM Transactions on Computer-Human Interaction, 7*(1), 29–58.

Abowd, G. D., Mynatt, E., & Rodden, T. (2002). The human experience. *IEEE Pervasive Computing-Journal on Mobile and Ubiquitous Systems, 1*(1), 48–57.

Ackermann, M. S., Hindus, D., Mainwaring, S. D., & Starr, B. (1997). Hanging on the wire: A field study of an audio-only media space. *ACM Transactions on Computer-Human Interaction 4*(1), 39–66.

Agamanolis, S. (2006). At the intersection of broadband and broadcasting: How ITV technologies can support human connectedness. *Proceedings of the European Conference on Interactive Television - EuroITV 2006*, 17–22.

Begole, J., Rosson, M. B., & Shaffer, C. A. (1999). Flexible collaboration transparency: Supporting worker independence in replicated application-sharing systems. *ACM Transactions on Computer-Human Interaction, 6*(6), 95–132.

Borghoff, U. M., & Schlichter, J. H. (2000). *Computer-supported cooperative work: Introduction to distributed applications*. Heidelberg, Germany: Springer-Verlag.

Bove, V. M., & Agamanolis, S. (2006). *Vision television*. Retrieved April 1, 2007, from http://web.media.mit.edu/~stefan/vtv/index.html

Coates, T. (2006). *Social software for set-top boxes*. Retrieved February 10, 2006, from http://www.plasticbag.org/archives/2005/03/social_software_for_settop_boxes/

Coppens, T., Vanparijs, F., & Handelkyn, K. (2005). *AmigoTV: A social TV experience through triple-play convergence*. Paris: Compagnie Financiere Alcatel.

Day, M., Aggarwal, S., Mohr, G., & Vincent, J. (2000). *Instant messaging / Presence protocol pequirements - RFC 2779.* The Internet Engineering Task Force. Retrieved April 1, 2007, from http://www.ietf.org/rfc/rfc2779.txt

Day, M., Rosenberg, J., & Sugano, H. (2000). *A model for presence and instant messaging (RFC 2778).* The Internet Engineering Task Force. Retrieved April 1, 2007, from http://www.ietf.org/rfc/rfc2778.txt, 2000b

Dewan, P., & Choudhary, R. (1991). Flexible user interface coupling in a collaborative system. In *Proceedings of the Conference on Human Factors in Computing Systems - CHI'91* (pp. 41–48). New York: ACM Press.

Dourish, P., & Bly, S. (1992). Portholes: Supporting awareness in a distributed work group. In *Proceedings of the Conference on Human Factors in Computing Systems - CHI'92* (pp. 541–547). New York: ACM Press.

Garfinkel, H. (1967). *Studies in ethnomethodology.* Englewood Cliffs, NJ: Prentice-Hall.

Gaver, W. W., Moran, T., MacLean, A., Lövstrand, L., Dourish, P., Carter, K. A., et al. (1992). Realising a video environment: EUROPARC's RAVE system. In *Proceedings of the Conference on Human Factors in Computing Systems - CHI'92* (pp. 27–35). New York: ACM Press.

Gross, T., Braun, S., & Krause, S. (2006). MatchBase: A development suite for efficient context-aware communication. In *Proceedings of the Fourteenth Euromicro Conference on Parallel, Distributed, and Network-Based Processing - PDP 2006* (pp. 308–315). Lost Alamitos, CA: IEEE Computer Society Press.

Gross, T., Egla, T., & Marquardt, N. (2006). Sens-ation: A service-oriented platform for developing sensor-based infrastructures. *International Journal of Internet Protocol Technology, 1*(3), 159–167.

Gross, T., & Marquardt, N. (2007). CollaborationBus: An editor for the easy configuration of ubiquitous computing environments. In *Proceedings of the Fifteenth Euromicro Conference on Parallel, Distributed, and Network-Based Processing - PDP 2007* (pp. 307–314). Los Alamitos, CA: IEEE Computer Society Press.

Gross, T., & Oemig, C. (2005). PRIMI: An open platform for the rapid and easy development of instant messaging infrastructures. In *Proceedings of the 31st EUROMICRO Conference on Software Engineering and Advanced Applications - SEAA 2005* (pp. 460–467). Los Alamitos, CA: IEEE Computer Society Press.

Gross, T., & Oemig, C. (2006). From PRIMI to PRIMIFaces: Technical concepts for selective information disclosure. In *Proceedings of the 32nd EUROMICRO Conference on Software Engineering and Advanced Applications - SEAA 2006* (pp. 480–487). Los Alamitos, CA: IEEE Computer Society Press.

Gross, T., & Prinz, W. (2004). Modelling shared contexts in cooperative environments: concept, implementation, and evaluation. *Computer Supported Cooperative Work: The Journal of Collaborative Computing, 13*(3–4), 283–303.

Gross, T., Stary, C., & Totter, A. (2005). User-centered awareness in computer-supported cooperative work-systems: Structured embedding of findings from social sciences. *International Journal of Human-Computer Interaction, 18*(3), 323–360.

Hammond, J., Gross, T., & Wesson, J. (Eds.). *Usability: Gaining a competitive edge*. Dordrecht, The Netherlands: Kluwer Academic.

Ignite Realtime. *Ignite Realtime: Smack API.* Retrieved May 1, 2007, from http://www.igniterealtime.org/projects/smack/index.jsp

ITU. *ITU Telecommunication Standardisation Sector (ITU-T)–IPTV Focus Group (FG IPTV).* Geneva, Switzerland: International Telecommunication Union. Retrieved February 10, 2006, from http://www.itu.int/ITU-T/IPTV/

Jones, Q., Grandhi, S. A., Terveen, L., & Whittaker, S. (2004). People-to-people-to-geographical-places: The P3 framework for location-based community systems. *Computer Supported Cooperative Work: The Journal of Collaborative Computing, 13*(3–4), 249–282.

Kirstein, P., & O'Hanlon, P. (2006). *Robust Audio Tool*. London: University College. Retrieved February 10, 2006, from http://www-mice.cs.ucl.ac.uk/multimedia/software/rat/index.html

Marca, D., & Bock, G. (Eds.). (1992). *Groupware: Software for computer-supported cooperative work.* Los Alamitos, CA: IEEE Computer Society Press.

MySQL. (2006). *MySQL: The world's most popular open source database.* Uppsala, Sweden. Retrieved February 10, 2006, from http://www.mysql.com/

NTL Group Ltd. (2006). ntl.com [Home page]. Available at http://www.ntl.com/

O'Brien, J., Rodden, T., Rouncefield, M., & Hughes, J. (1999). At home with the technology: An ethnographic study of a set-top-box trial. *ACM Transactions on Computer-Human Interaction, 6*(3), 282–308.

Oehlberg, L., Ducheneaut, N., Thomton, J. D., Moore, G. E., & Nickell, E. (2006). Social TV: Designing for distributed, sociable television viewing. *Proceedings of the European Conference on Interactive Television - EuroITV 2006*, 251–262.

Preece, J. (2000). *Online communities: Designing usability, supporting sociability*. New York: Wiley.

Regan, T., & Todd, I. (2004). Media center buddies: Instant messaging around a media center. In *Proceedings of the Fourth Nordic Conference on Human-Computer Interaction-NordiCHI 2004* (pp. 141–144). New York: ACM Press.

Resnick, P., & Varian, H. R. (1997). Introduction to special issue on Recommender Systems. *Communications of the ACM, 40*(3),56–58.

Rheingold, H. (1996). *The virtual community*. Retrieved November 20, 1996, from http://www.rheingold.com/vc/book/

Rodden, T., & Blair, G. (1991). CSCW and distributed systems: The problem of control. In *Proceedings of the Second European Conference on Computer-Supported Cooperative Work - ECSCW'91* (pp. 49–64). Dordrecht, The Netherlands: Kluwer Academic.

Sanches, J. A., Leggett, J. J., & Schnase, J. L. (1994). HyperActive: Extending an open hypermedia architecture to support agency. *ACM Transactions on Computer-Human Interaction, 1*(4), 357–382.

Schulzrinne, H., Casner, S., Frederick, R., & Jacobson, V. (1993). *RTP: A transport protocol for real-time applications - RFC 1889.* The Internet Engineering Task Force. Retrieved February 10, 2006, from http://www.ietf.org/rfc/rfc1889.txt

Sun Microsystems Inc. (2006). *Java technology*. Sun Microsystems Inc., Retrieved February 10, 2006, from http://java.sun.com/

The Apache Software Foundation. *WebServices - Axis architecture guide.* World Wide Web Consortium. Retrieved February 10, 2006, from http://ws.apache.org/axis/java/architecture-guide.html

Weiser, M. (1991). The computer of the 21st century. *Scientific American, 9,* 94–104.

Weiser, M. (1993). Some computer science issues in ubiquitous computing. *Communications of the ACM, 36*(7), 75–84.

Winer, D. (1999). *XML-RPC specification*. Retrieved February 10, 2006, from http://www.xmlrpc.com/spec

YouTube. (2006). *YouTube—Broadcast yourself*. Retrieved February 10, 2006, from http://www.youtube.com/