together anywhere, together anytime

ICT-214793

TA2

Together Anywhere, Together Anytime

Large Scale Integrating Project

ICT – Networked Media

D3.5

Summary Report - Application Design and Implementation

Due date of deliverable: 30 September 2011

Actual submission date: 16 November 2011

Start date of project: 1 February 2008 Duration: 48 months

Lead contractor for this deliverable: BT

Final Version of 15 November 2011

Confidentiality status: Public (Not yet approved by the European Commission)

Abstract

This document provides a public record of some of the work carried out in the EU funded project TA2 "Together Anywhere, Together Anytime" (TA2). The project is funded under the EU 7th Framework programme and seeks to understand how technology can help make communication and engagement easier between groups of people separated in time and space. The project has designed and built a number of concept demonstrators to help explore this question. This document provides a description of the concept demonstrators used within the project; the motivation for using demonstrators; and the main findings from the evaluations carried out with each demonstrator.

The five key concept demonstrators are:

- *Family Game* – bringing the experience of a family board game to people in separated households,
- *Music Tuition* – a system testing the value of multiple camera set ups in remote music tuition,
- *Storytelling* – a means for allowing bedtime stories to be shared between different households,
- *MyVideos* – a system to allow the development of personalised videos collated from content captured at a school concert by many of the audience,
- *Connected Lobby* - a means by which households engage in any of the above activities based on presence technologies displayed through a TV.

The document illustrates how each of the demonstrators offers different insights that are useful to the project, either because the demonstrators exercise different technological capabilities, or different end user characteristics, or different demonstrator characteristics (e.g. whether the interaction involved is synchronous or not, or whether the experience will be evaluated in a lab or within the environments that it is expected to be used).

Brief descriptions of the evaluations and the findings form these evaluations are provided.

Target audience

The document is designed to be read by members of the public outside the project who may be interested in how technology may be used in the future and in how the EU is funding research to better understand this field.

Disclaimer:

This document contains material, which is the copyright of certain TA2 consortium parties, and may not be reproduced or copied without permission. All TA2 consortium parties have agreed to full publication of this document. The commercial use of any information contained in this document may require a license from the proprietor of that information.

Neither the TA2 consortium as a whole, nor a certain party of the TA2 consortium, can warrant that the information contained in this document is capable of use, nor can they warrant that use of the information is free from risk. They accept no liability for loss or damage suffered by any person using this information.

This document does not represent the opinion of the European Community, and the European Community is not responsible for any use that might be made of its content.

Impressum

Full project title: Together Anywhere, Together Anytime
Title of the work package: WP3 – Application Design and Implementation
Document title: D3.5 Summary report – Application Design and Implementation
Editor: BT
Work package Leader: BT
Project Co-ordinator: Peter Stollenmayer, Eurescom
Technical Project Leader: Doug Williams, BT
This project is co-funded by the European Union through the ICT programme under FP7

Copyright notice

Executive Summary

This document provides a public record of some of the work carried out in the EU funded project TA2 "Together Anywhere, Together Anytime" (TA2). The project is funded under the EU 7th Framework ICT programme and seeks to understand 'how technology can help make communication and engagement easier between groups of people separated in time and space'. The project has designed and built a number of concept demonstrators to help explore the questions and this document provides a description of the demonstrators used within the project; the motivation for using demonstrators; and the main findings from the evaluations carried out with each demonstrator.

The five concept demonstrators researched during the four years of the project and refined by use during the final eighteen months are described. These are:

- *Family Game* – bringing the experience of a family board game to people in separated households,
- *Storytelling* – a means for allowing bedtime stories to be shared between different households,
- *MyVideos* – a system to allow the development of personalised videos collated from content captured at a school concert by many of the audience,
- *Connected Lobby* - a means by which households engage in any of the above activities based on presence technologies displayed through a TV.
- *Music Tuition* – a system testing the value of multiple camera set ups in remote music tuition.

The TA2 concept demonstrators were developed through a user centred design method; a process in which the needs, wants, and limitations of end users are given extensive attention at each stage of the design process. For each demonstrator the following is discussed:

- The value that each demonstrator offers the project
- A short narrative of the development of the demonstrator
- Some details about the evaluations carried out and the results yielded from them
- A brief anticipation of how each demonstrator is likely to develop

The observations made are summarised as follows:

Family Game:

- Whilst game playing is an important social activity for groups and does help to build relationships, gamers report that tactile elements of the game (the board, the pieces etc.) are important aspects that should, if possible, be retained.
- Game design should focus on game play in which communication is required; cooperative games are likely to be a good choice in this regard.
- In order to enhance the value derived from the communication, the game and video elements should be composited on the same screen to encourage eye contact. This requirement will emphasise capabilities of the visual composition engine component developed in TA2 and are being tested in our trials scheduled for November 2011.

Storytelling:

- The ability to perform evaluations within people's homes provides insights into the domestication of technology that cannot be achieved in the laboratory and that cannot be easily anticipated. Such experiments, whilst a compromise in terms of assessing 'bleeding edge' technology, remain extremely valuable.
- Quantitative data on the number of times books are read, the time spent on each page whether books are re read and whether participants utilise the interactive elements is being collected and will be available in late 2011/early 2012
- Users have invented their own use case for shared applications which require flexibility in the placement of microphones and cameras in order to facilitate a particular kind of sharing (a view on a crossword puzzle) or to improve the fidelity of the interaction (by moving the microphone closer to the speaker). Future designs of such system should explore the use of mobile video and capture devices and to consider whether current models for spatialising audio and orchestrating the video are robust enough to accommodate capture devices that can move.
- All users without exception have asked whether they can keep the TA2 setup. This is an encouraging qualitative indicator of the value users perceive in the technology.

MyVideos:

- Users believe that systems like MyVideos would encourage them to capture more video material and to share more video. Assuming that the sharing of stories is an important part of building togetherness this would suggest that the ability to generate personalised stories from shared content will help nurture relationships between people who know each other well.
- Users of the early *MyVideos* system were (generally) pleased with the personalised presentations that they produced with the system.

Connected Lobby:

- The television is seen as an excellent device for group communication with associations to social and relaxation oriented behaviour.
- Video communications through the TV is seen as an addition to, and not a replacement for, telephone calls. One user referred to it as being more of an event – "like going for a cup of coffee". This perception is useful in understanding how users will incorporate Video Communications into their lives and therefore an important insight to be used in presenting, through marketing message for instance, how such products could be used.

Music Tuition:

- The system developed by TA2 enabled music lessons that according to the separated pupils and tutors are effective and, according to some comments, "just like a normal lesson".
- The use of multiple camera views appeared important for some instrument types but not all.
- Ethnographic observations of real face to face lessons highlight the importance of the physical music score during lessons. We anticipate the systems would be more useful if an effective way of sharing the paper based annotated score, as used by the pupil, could be devised.

- The standard audio capture provided by the Polycom system used in the experiments lacked the dynamic range to effectively encode loud piano and horn pieces; improved audio capture based on analogue and mixing desk like capabilities will be an essential component of improved music tuition systems.

These are demonstrator-specific insights that are useful for the evolution of each demonstrator. Overall it is clear that a user-centred design approach to the development and evaluation of demonstrators has helped us to understand how the demonstrators affect the nurturing of social relationships between people who know each other well. Some of the conclusions made can be related to the technology capabilities developed within the project. *Music Tuition* and *Storytelling* both revealed users demands for roving cameras that could capture additional shared content (a musical score and a crossword puzzle were the particular examples). We can infer from this that the ability to intelligently compose on a shared screen, images involving both the view of the remote person and the additional camera, are valuable. The ability to capture, transmit and intelligently compose multiple views is one of the key developments within TA2.

List of Authors

Breitung, Michael – Fraunhofer IIS

Bulterman, Dick – CWI

Coezijn, Etienne – Philips

Färber, Nikolaus – Fraunhofer IIS

Frey, Alex – Limbic Entertainment

Kegel, Ian – BT

Ljungstrand, Peter – Interactive Institute

Spitzner, Christian – Fraunhofer IIS

Türck, Clemens – Ravensburger

Williams, Doug – BT

Table of contents

List of figures

Abbreviations

AS3	ActionScript 3. A coding language used in developing applications in Flash
CE-HTML	HTML for Consumer Electronics
CM	Communication Manager. A component in the TA2 architecture
CM-LC	Communication Manager Local Control. A component in the TA2 architecture
CM-SC	Communication Manager Session Control. A component in the TA2 architecture
CMC	Computer-Mediated Communication. Any communicative transaction that occurs through the use of two or more networked computers.
HTML	Hypertext Markup Language the predominant mark-up language for web pages
HTTP	Hypertext Transfer Protocol a networking protocol for distributed, collaborative, hypermedia information systems it is the foundation of data communication for the World Wide Web
IPTV	Internet Protocol Television. A managed TV service delivered by encapsulating the video as data transmitted using protocols normally used in the Internet.
LAMP	Linux (operating system), Apache HTTP Server, MySQL database software and PHP (server-side scripting language)
LED	Light-Emitting Diode
MUC	Multi-User Chat
RFID	Radio Frequency Identification: a technology that uses radio waves to transfer data from an electronic tag.
RTMP	Real Time Messaging Protocol a TCP-based protocol which maintains persistent connections and allows low-latency communication
SDS	Semantic Differential Scaling a method of a rating subjective, cultural and/or emotional association that some word or phrase carries, in addition to the word's or phrase's explicit or literal meaning.
SIP	Session Initiation Protocol – a signalling protocol widely used for controlling communications sessions such as voice and video calls over Internet Protocol (IP). The protocol can be used for creating, modifying and terminating two-party or multiparty sessions
SMIL	Synchronous Multimedia Integration Language is used to describe multimedia presentations. It defines mark-up for timing, layout, animations, visual transitions, and media embedding, among other things.
TA2	Together Anywhere, Together Anytime (this project)
TUI	Tangible User Interface
UI	User Interface

VoIP	Voice over Internet Protocol
VCE	Visual Composition Engine, a TA2 developed software component that renders a wide range of visual forms, including video streams, Flash files and text and graphic overlays using SMIL
XML	Extensible Mark-up Language - a set of rules for encoding documents in a machine readable form
XML-RPC	Remote Procedure Calling a remote procedure call (RPC) protocol which uses XML to encode its calls and HTTP as a transport mechanism.
XMPP	Extensible Messaging and Presence Protocol; an open standard communications protocol for message oriented middleware based on XML for near real-time extensible instant messaging (IM), presence information and contact list maintenance http://en.wikipedia.org/wiki/Presence_information.

1 Introduction

The research project *"Together Anywhere, Together Anytime"* (TA2) seeks to understand how technology can "enhance relationships between groups of people who are separated in space and time".

TA2 has adopted a user centred design approach to answer this question by which we mean we have given extensive attention to the needs, wants, and limitations of end users at each stage of the design process. We have defined a number of use cases and then designed and built the technology to embody each use case in a concept demonstrator. The concept demonstrators are described in this document.

As part of the TA2 project we have assimilated understanding of group based social behaviour and also of computer mediated communications. We recognise that both synchronous and asynchronous activities can have an impact on the way relationships are nurtured. A particular focus of the project is communications between multiple locations (more than two) and between groups of people, rather than just between individuals. The requirements and technology challenges associated with this focus have been discussed elsewhere (Williams, 2011) but they include:

- **A shared focus**: an activity which is common to all ends of the experience that provides a common fun shared activity that aids in the building of social bonds between the participants.
- **HD video communications**: so that participants can clearly see each other and so that they gain both peripheral awareness of people at each end and if necessary, some indication of eye contact and the ability to transmit and interpret gesture and body language
- **HD audio**: (at least stereo) audio transmitted using super wide band audio bandwidth (>7kHz) in order to support relaxed free space natural communications between participants at each location
- **Multiple HD cameras**: in order to capture different views of each end
- **Orchestration**: a means of selecting the camera view that provides the 'best' representation of the remote ends to all parties involved in the shared experience

This document describes the way that TA2 has used demonstrators to answer its high level research question and describes our findings to date. It describes, in Section 1.1, why demonstrators are used, with reference to the project's goals and with some reflections about the perils (and value) of using them. Section 1.2 explains how the particular demonstrators used in TA2 were chosen. Sections 2 to 5 provide more detail on each of the five demonstrators which are:

- *Family Game* – which brings the experience of a family board game to people in separated households.
- *Storytelling* – a means for allowing a bedtime story to be shared between different households.
- *My Videos* – a system to allow the development of personalised videos collated from content captured at a school concert by many of the audience.
- The *Connected Lobby* - a means by which households engage in any of the above activities based on presence technologies displayed through a TV.
- *Music Tuition* – a system testing the value of multiple camera set ups in remote music tuition.

For each demonstrator the following is discussed:

- The value that each demonstrator offers the project
- A short narrative of the development of the demonstrators
- Some details about the evaluations carried out and the results yielded from them

- A brief anticipation of how the demonstrators is likely to develop

A brief summary and reflection on the use of demonstrators is offered through section 7.

1.1 The concept demonstrators

The project recognises that its goal of helping to 'nurture social relationships' is a high level one that is not easily correlated with hard technical experimental data that can be measured in a laboratory. For this reason we chose to explore the deployment of technology capabilities within a number of concept demonstrators, with which people could interact at different levels and through which we hoped to elicit insights into the way technology can be harnessed to nurture social relationships between family and close friends.

We believe that this is the most obvious way of demonstrating and assessing the value of the technology capabilities in contexts most likely to have an inferred impact on the relationships enjoyed between family and friends.

Performing rigorous and definitive evaluations in this way is not easy. Particular challenges that this approach introduces include:

- Assessing the unique impact of an individual technological component within a complex system.
  The technologies used within TA2 are deployed together with other technologies and the apparent experience of one component (which may be excellent) could be adversely affected by the poor performance of another.
- Assessing the impact of an individual technology component when its impact is intrinsically bound to activity within which it is deployed.
  The best technology deployed within an application which does not benefit from the capabilities offered by the technology will have no impact on relationships.
- Assessing the impact of the technology within experiences that will be assessed subjectively by a limited number of people.
  The best technology deployed in an excellently designed application that provides an experience that is unappealing for the subjects of the experiment, will have no positive impact on their relationships. It does not mean that the technology is of no use nor that its impact will not be appreciated by some people.
- Designing and building activities whose impact on relationships could be evaluated within the life of the project.
  Excellent technologies in excellent applications that cannot be reliably or frequently used in people's homes will not provide meaningful assessments of the impact on relationships.

To navigate these challenges TA2 chose to design and build a number of different demonstrators, which were designed to use different constellations of technology, to appeal to different sorts of people, and to provide different sorts of evaluation opportunities.

Having chosen to test our research question using a number of different demonstrators, we have to get three things right.

- The deployment of the capability
- The choice of the demonstrator
- The evaluation; the objective and method of the evaluation both need to be correct.

These are not independent:

- Different levels of *capability* offered by a technology can affect the way the designer chooses to build the demonstrator and hence change its nature.
- Different levels of technology *stability* affect the forms of evaluation that can be considered. Stable technologies can be evaluated in long term studies; fragile configurations can only be assessed in lab based evaluations – and these differences will change what you hope to evaluate.
- And of course, the evaluation objective may affect the nature of the technology you try to develop. If you are convinced that audio quality is an important variable then you will choose demonstrators where you believe audio, above all other considerations, is key.

So, by design and necessity, the project is highly iterative.

1.2 Rationale for the choice of concept demonstrators

In this document we report on the five concept demonstrators under evaluation during the final 18 months of the project.

The core methodology used to choose the different demonstrator was a user centred design. This method was chosen as it should ensure, as far as it is possible to do so, that the capabilities we develop are deployed in activities designed from the users' perspective. We started by developing clear pictures of the people for whom we were designing applications which were described in personae and, for the family groups, familia. These provided rich pictures about our users including their lifestyles and of the relationships that existed between them. Through brainstorming we generated a long list of demonstrators ideas that was subsequently refined, based on the extent to which we believed the demonstrators would benefit from the capabilities the project was intending to develop. This refinement has continued through the project causing the nature of the demonstrators to adapt during the project. This process provided us with a range of different demonstrators, all designed through a user centred methodology, that together were intended to enable sensible evaluations to be made of the technology capabilities we were developing and on the impact they may have on nurturing *social relationships between family and close friends.*

Figure 1 shows what each of the five demonstrators brings the project. A number of different coordinates can be considered, such as:

- The design/technology aspect of the capabilities that can be evaluated
    - application: design and user interface concepts
    - video: the capture, encoding, transmission and representation
    - audio: the capture, encoding, transmission and play-out
    - analysis: the effectiveness of automatic audio and video analysis in delivering meaningful cues for orchestration
    - presence: the way availability information is managed for groups on shared interfaces
    - orchestration: the automatic selection of camera view and composition based on rules driven from low level audio and video analysis
    - communications manager: the way different devices can communicate and be managed

- The nature of the demonstrators
    - Two or more than two end points?
    - Synchronous, near synchronous or clearly asynchronous interactions?
    - For individuals or for groups?
    - Stable enough for longitudinal end user evaluations, or only suitable for lab based evaluations?
- The age of, and the nature of the relationships between, the participants.

Figure 1 shows that each of the demonstrators allows the project to probe different elements along different coordinates and that together just about all possible coordinates are explored. This illustrates the value of exploring more than one demonstrator. In addition of course the decision to choose multiple demonstrators helped to illustrate how components could be re-used with in different contexts.

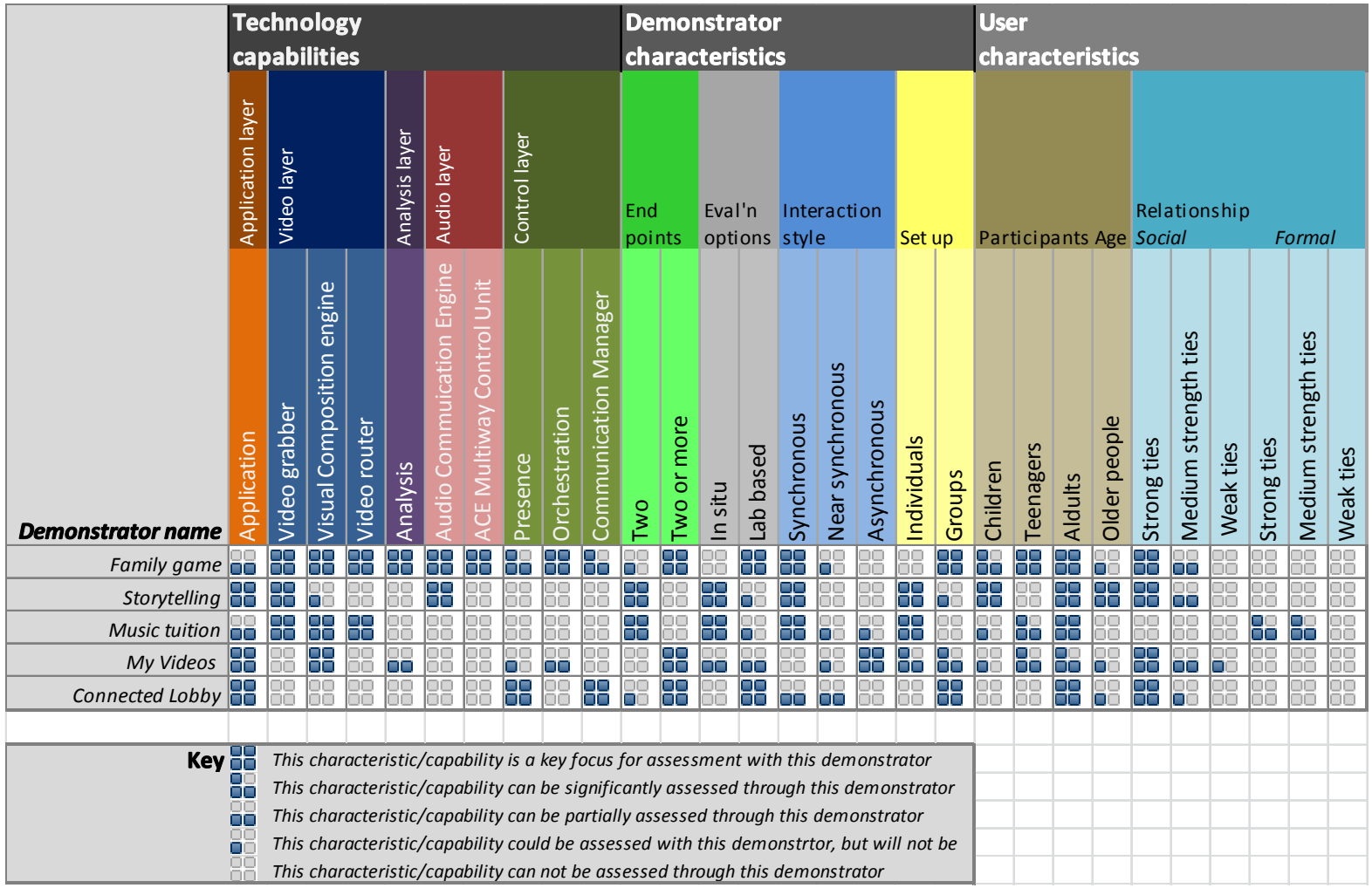


Figure 1 Illustrating that the TA2 demonstrators together provide a way of evaluating the widest possible gamut of technology capabilities, demonstrators characteristics and user characteristics

2 Family Game

Description

The *Family Game* demonstrator brings the experience of a casual family game (such as a board game), an experience which normally relies very much on people's interaction within one room, to separated households. The essentials of a family game are interaction, communication, sharing a common experience and above all having fun together.

Technology

The *Family Game* uses all components in the TA2 architecture. It is designed to be played between three locations, uses multiple cameras at each location and requires the Orchestration Engine and the Visual Composition Engine (VCE). The Orchestration Engine is a network-based component that chooses, in real time, the 'best' audio and video presentation to show at each TA2 location. Decisions about what to show are based on analysis of the audio and video signals together with contextual information about the interaction. The VCE composes visual images at each TA2 location. It renders real-time video in combination with pre-recorded media from a local repository. It can also incorporate other forms of content, such as Adobe Flash, from the game.

Social Science

Playing a board game together is a valuable way of fostering strong ties. Board games are often immersive and emotional experiences that create shared memories and thus enhance togetherness in the long term. They can be regarded as *interaction rituals* which are important for maintaining social relationships and for building social cohesion and social identity.

Evaluation

Family Game is being assessed through lab tests in which the qualitative responses of people playing a traditional board game in a single room are compared with those of triallists evaluating the TA2 *Family Game* played between three locations.

Learning

Game design should focus on game play in which communication is required; cooperative games are likely to be a good choice in this regard.

Tactile elements of the board game (the pieces etc.) should be retained.

In order to enhance the value derived from subtle nuanced communication common in social game play, the game and video elements should be composited on the same screen to encourage eye contact. This will emphasise capabilities of the visual composition engine developed in TA2.

The *Family Game* demonstrator aims to bring the experience of a family (board) game, an experience which normally relies very much on people's interaction within one room, to separated households. The essentials of a family game are interaction, communication, sharing a common experience and above all having fun together.

The TA2 *Family Game* has a space theme. The players are members of one space ship. This space ship is stranded in space – far from Earth and it has little fuel left. The goal for the players is to gather enough fuel to return home safely. They can do this by using the remaining fuel to reach the next planet or asteroid on their way. There they can gather sufficient fuel to jump to the next planet; and so on, until they are close enough to Earth. All the players together have to decide which planet to head for, calculating risks and chances on each planet.

The visual aspects of the game are composited and rendered on the TV screen. Some of the chance aspects in the game are determined using playing cards which include RFID tags that can be read using a reader provided at each location to provide input to the overall system. Tasks, like gathering fuel on the planets, are achieved through effectively completing mini-games which use the body as input device.

The game is cooperative; all players share the same goal, they all succeed or lose together. Lately, there have been several board games on the market which share this cooperative approach and have been highly successful and got good reviews including: Wer war's?; Space Alert; Pandemic and Shadows over Camelot.

2.1 *Family Game*: the value of this demonstrator

Family Game enables, as can be seen in Figure 1, a use case that involves just about all of the technical components notably including the analysis and orchestration components and enables the project to test whether these functions can be made to work as intended and to begin to explore whether the effect they create is appealing. This will be evaluated through lab based evaluations as the complete technical system comprises many computers and cannot be realistically assembled with user's homes.

The *Family Game* demonstrator offers valuable insights and research possibilities in the following areas:

Social Science:

Enhancing relationships of families and friends is the central theme of TA2.

Playing a (board) game **together** is a valuable way of fostering strong ties. This is typically an immersive and emotional experience which creates shared memories and thus enhances togetherness in the long term. It can be regarded as an *interaction ritual*. Such interaction rituals are important for maintaining social relationships and for building social cohesion and social identity (Durkheim, 1971). Others ((Goffman, 1967), (Collins, 2005) and (Ling, 2008)) have noted that an important part of interaction rituals is performing mutual activities; it has also been suggested ((Kock, 2004), (Collins, 2005)) that interaction rituals making use of the full expressive capacity of human beings will make a stronger impact than those only based on language. We believe that TA2 sets out to achieve a form of communication that is much less limited, in terms of its embodied interaction, than forms of mediated communication that are popularized today (Dourish, 2001). In addition TA2 seeks to provide a focus on groups and not individuals as actors ((Ljungstrand, 2008). We believe that the focus on a collaborative board game is therefore an excellent choice of a shared activity that can help nurture the relationships between groups.

We want to compare people's experience of playing a casual social game over the TA2 system with people's experience of playing a casual social game in one room.

Technical Development:

As can be seen in Figure 1 and in Figure 2, the *Family Game* makes use of all the TA2 technical components and is furthermore the only demonstrator to connect synchronously systems in three distant locations over standard broadband internet. It is, technologically, the most demanding demonstrator and because of this it is only possible to evaluate *Family Game* in a lab setting. In situ trials are unrealistic due to the large number of computers involved in delivering *Family Game*.

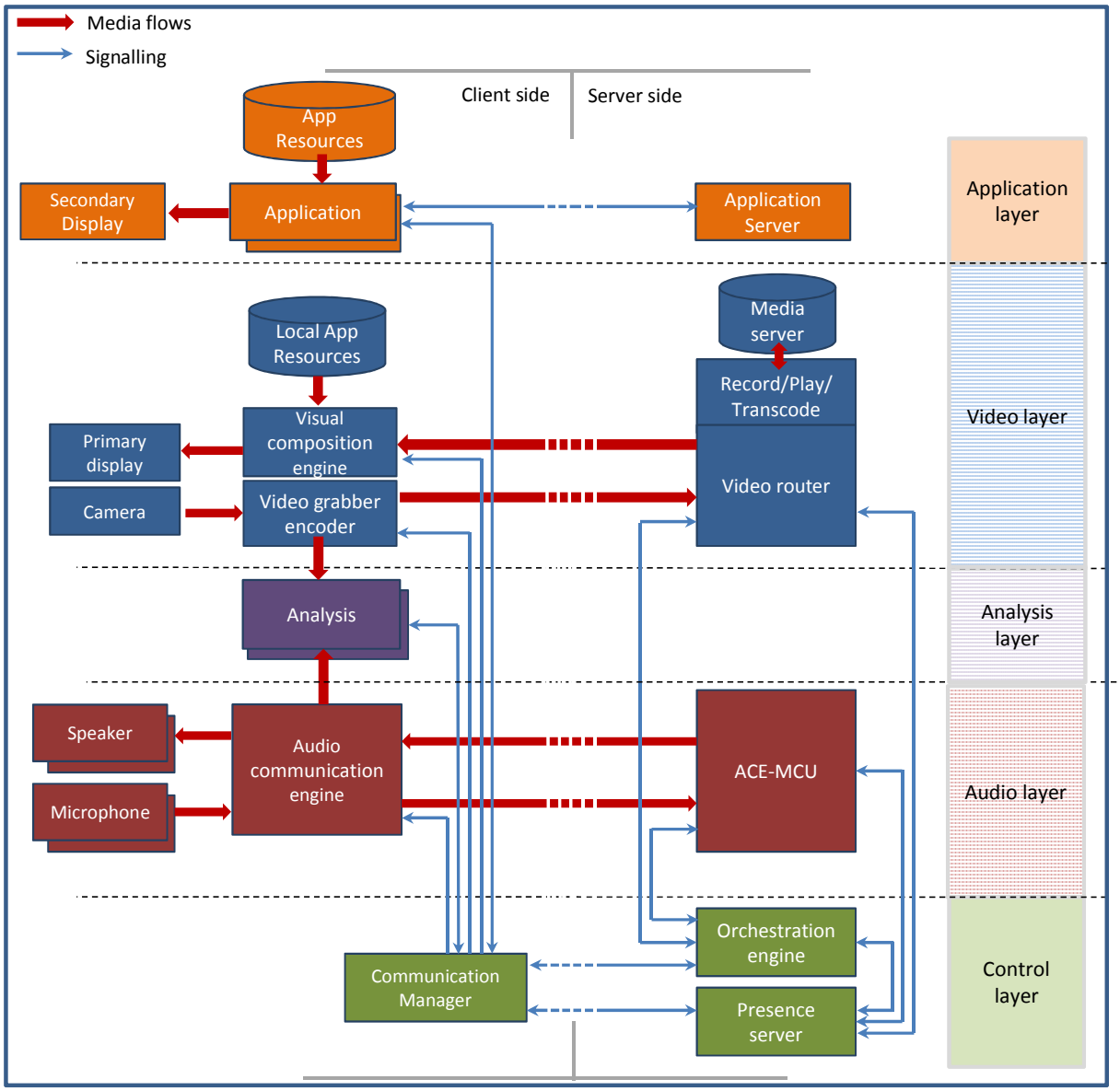


Figure 2 The technical architecture used to deliver the *Family Game* demonstrator

This technical complexity is also the reason why this demonstrator is available only in the later stage of the project. This leaves less time to evaluate the whole experience with end users. Therefore additional evaluations have been carried out looking at parts of the experience including experiments on: audio quality and its impact on user experience; the impact of manual, automatic or no orchestration on user experience; the playing of a board game over video conference system.

Exploitation potential:

Gaming is a strong driver for platform and infrastructure (e.g. network) development for many interactive systems (see development of Personal Computers, Mobile Phones, gaming consoles).

Gaming Industry partners of the consortium like Ravensburger and Limbic are highly interested in which forms of gaming will develop in the future out of new hardware systems such as those demonstrated within TA2.

Infrastructure developers like Alcatel and BT or hardware manufacturers like Philips are highly interested in learning which new applications are strong enough to push the interest in new platforms and services.

2.2 *Family Game:* concept development

In developing The *Family Game* we sought a game mechanism which relied on communication and interaction between the players, in order to create suitable challenges to the TA2-System.

The space travel topic, where people might be separated in space within closed environments, connected only through video and audio was found to be well suited to the TA2 environment.

Inspired by the award winning cooperative board game "Space Alert", published in autumn 2008, we developed a TA2 version in 2009 with a similar game play. Players are bound together on a mission, flying a space ship into dangerous regions and having to avoid damage through attacking enemies. The players need to work together under time constraints to pump energy to the right parts of the ship, to avoid being damaged by enemy fire and to shoot and destroy enemies and asteroids on collision course.

At first the game setup was tested offline (with game board materials) with game experts and friendly users. Then the game was implemented in Flash. The whole TA2 system with initial implementation of Analysis and Orchestration was set-up for the first time in early 2010. There were two separated living rooms equipped with TA2 systems. On this system we could make first tests under "real-live" conditions. User groups could play the game and their experiences could be evaluated.

From this test run, we learned many lessons (see section 2.3), related to the game mechanism, the user interface and the use of technology.

With these findings in mind, we redesigned the basic game play and user interface as described above.

In meetings at Gothenburg (Interactive Institute) in June 2010 and Antwerp (Alcatel) in July 2010 this game design was discussed and improved. An Evaluation process was set up in Delft in August 2010 and the new game structure was finalized in September 2010 at a meeting at Goldsmiths, London.

Work up to September 2011, included implementing the new und updated elements for the system and integration. The tasks which filled the better part of 2011 were:

- Implement the new game mechanics in the Flash game
- Create new user interface (overlaid over video stream)
- Redesign RFID-reader for the new game mechanics
- Create three mini-games together with Analysis
- Define communication between Game Engine, CM, orchestration, and VCE
- Create orchestration concept for three location orchestration
- Make the system work between three locations
- Implement the mini-games

Since it was clear that the whole system will only be running quite late in the project, smaller evaluations of the single parts had been done continuously while developing.

For example, multi-party session control has been tested during a workshop in Erlangen, December 2010. The Three-way Orchestration has been tested in Graz February and March 2011 and in three technical integration workshops until September 2011, the components have been brought together.

Figure 3 The Fraunhofer demonstration set up including a table top display and large screen video used to test the audio chain in the playing of a game remotely

By mid October 2011, the components are, more or less, integrated but some tweaks and improvements are still being made. Test runs with friendly users are planned in October and evaluations with external triallists are planned in November 2011.

2.3 *Family Game:* Evaluation results

After the first game design was defined but before any of the system components were developed, a focus group was organized in 2009 in order to let potential users evaluate the Family Game.

The focus group was organized in cooperation with *Spelgroep Phoenix* (playing group), see Figure 4, a not-for-profit foundation that enables its members to come together one night a week to play board games. The focus group was conducted during one of those nights, at their location. This approach allowed us to study a group of people in a natural setting.

Figure 4 Spelgroep Phoenix, the group who helped evaluate the game even before any technical components had been built

The users were presented the *Family Game* in a board game version together with the description of the concept of a TA2-system and how this game could be played over the TA2-system.

The various remarks from the focus group are summarized below. A more detailed version can be found in the Deliverable D8.4 Evaluations of TA2 concepts (available on the TA2 Website http://www.ta2-project.eu/deliverables/deliverables.html).

Overall comments:

- Group dynamics and social interactions between the people playing are key to the board game play experience and to the experience of togetherness: playing with the game; playing between the players; the game *on* the table and the game *over* the table.
- Even with the TA2 system, there is still the need to arrange meetings and being at a certain place at a certain time. This needs to be considered within the whole system.

Considerations:

- One idea from the focus group, was to create a set-up in which the screen and the table are very close to each other, so that the screens disappear, and the tables grow into one, shared table, in order to facilitate the experience of togetherness.
- A kind of video and audio communication is required that enable people to experience togetherness between *here* and *there*, and to communicate with each other in very subtle ways.
- The touch and feel of a board game and of the physical pieces is important and needs to be retained even if you are playing online, in order to facilitate a good game play.
- Position the application as a platform, as a console, rather than as a stand-alone application. Many other games can be played on this TA2 platform. The platform is different from current platforms or consoles, because it uses a multi-touch, multi-user table and tactile elements—which is different from, e.g. Nintendo Wii or Microsoft Xbox.

The first test run of the whole system was done in early 2010 in Antwerp, at Alcatel's premises. It showed that the (then available) components of the system worked together quite well. But it also showed that the separated interface (one table screen and one wall screen) did not support communication well. Players tended to look at the screen on the table in front of them most (possibly because of the perceived complexity of the way the game was represented on that screen). Also there was some criticism from the testers that the game involved shooting elements and killing evil alien space ships.

These evaluation findings led to the following redesign principles, which then led to the current game design and setup:

- Reduce complexity to broaden the target group for evaluation
- Focus the game onto the wall screen (only small input units on the table) to enhance communication
- Get rid of the "killing" elements
- Introduce mini-games which make use of Video and Audio Analysis

Figure 5 The final setup of one demo-living-room at Alcatel

2.4 *Family Game:* What next?

The immediate next steps are to improve the stability of the system and to integrate all the components related to the game and the supporting technology. When this has been achieved user tests and evaluations will be conducted.

Evaluation topics and plans will be described in detail in the deliverable D8.8 Evaluations of TA2 concepts - version 3 (available on the TA2 Website http://www.ta2-project.eu/deliverables/deliverables.html).

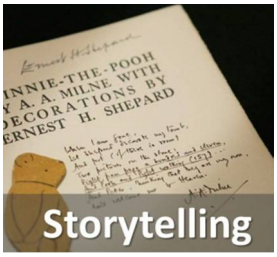
Commercial exploitation possibilities of the TA2 system and especially the *Family Game* will be discussed and will be reported in the Deliverable D8.9 Business model feasibility assessment (becoming available on the TA2 Website http://www.ta2-project.eu/deliverables/deliverables.html).

The Games companies have also learned about the challenge of building board games for parties in different locations; whilst the learning includes mostly tacit, difficult to codify knowledge, it relates to an improved understanding of how games need to be designed to fit the separation in space and time. One issue which we now understand needs careful consideration is managing (through design) what different participants can see and manipulate together.

The technology and infrastructure providers also gained tacit knowledge on what is important for the content providers, how they prefer to work and at which stages of the development process they should be included.

With the knowledge (learned from technology integration and evaluation results) and soft skills (gained through the highly networked cooperation) combined, there are good chances that parts of the TA2-system or derivations thereof which use gaming elements will come to market in some way or another.

3 *Storytelling*



Description

The S*torytelling* demonstrator attempts to enable adults and children to share something of the magical intimacy of reading a book together even when they cannot be together. The demonstrator uses iPads (which host the storybooks) and a PC connected to the TV which enables the adult and child to see each other in HD video quality. It also provides special effects including: animations on the iPad; overlay graphics on the TV; and sounds that can be triggered by clicking on objects on the page displayed in the iPad.

Technology

Storytelling uses consumer grade USB cameras and stereo microphones to allow HD video in 720p25 resolution and 48 kHz audio bandwidth. The total bit rate for H.264 video and AAC-ELD audio is 2 Mbps and all components are implemented in software on a Linux PC with small form factor and low noise. A dedicated iPad app synchronizes the book pages on both ends and communicates with the PC to trigger sound effects and graphic overlays. For the latter a simplified visual composition engine is used. In order to achieve the robustness and usability required for real home environments, the system is based on a light weight version of the full TA2 architecture.

Social Science

Bedtime stories encourage children to read and to love books, but they also provide precious time in one another's company. Many special memories can be created and, as well as paying special attention to the story, lessons and values can be learned based on the characters and the plot. The regular shared experiences, when reading a chapter each night, support social cohesion through interaction rituals. Enabling *Storytelling* will help nurture relationships.

Evaluation

Evaluation is conducted through in situ trials within people's homes. Systems are being made available along with a number of books, and both qualitative responses and quantitative usage data will be collated. About 10 families will be involved with the evaluations each using the system for between 4 and 6 weeks

Learning

All those who have used the system want to keep it; but discussion of the way they use and try to use the system has led to some adaptations such as a roving camera to support alternative uses of the system such as working on crosswords together.

The in-situ trials offer perhaps the strongest endorsement we have yet received that the ability to see each other in HD quality whilst engaging in a shared activity is precious and that it helps to nurture relationships.

TA2 decided to target the intimate activity of *Storytelling* as a concept demonstrator. The demonstrator was designed to allow stories to be told in an engaging even when the storyteller (a grandparent perhaps) and the child are not together.

The challenge is to reduce the impact of spatial separation such that both participants experience the same feelings that they would enjoy if they were in the same room. At the same time we also try to enhance the overall *Storytelling* experience by using multimedia extensions (sound, images, animation).

Storytelling uses a PC for HD-Videoconferencing as well an iPad-app for the interactive storybook. Hence, the participants will see and hear each other in high quality via the TV while typically sitting on the couch in the living room. Each side is holding an iPad which is running the storybook application. This application is basically a book reader with the special feature that both apps are synchronized, i.e. if one iPad turns the page, the page is also turned on the other iPad. Special effects including: animations on the iPad; static overlay graphics on the TV; and sounds that are played out in high quality (and volume) via the TV or HiFi sound system can be triggered by clicking on objects on the page displayed in the iPad. The basic scenario of the application, illustrated in Figure 6, was devised by Ravensburger.

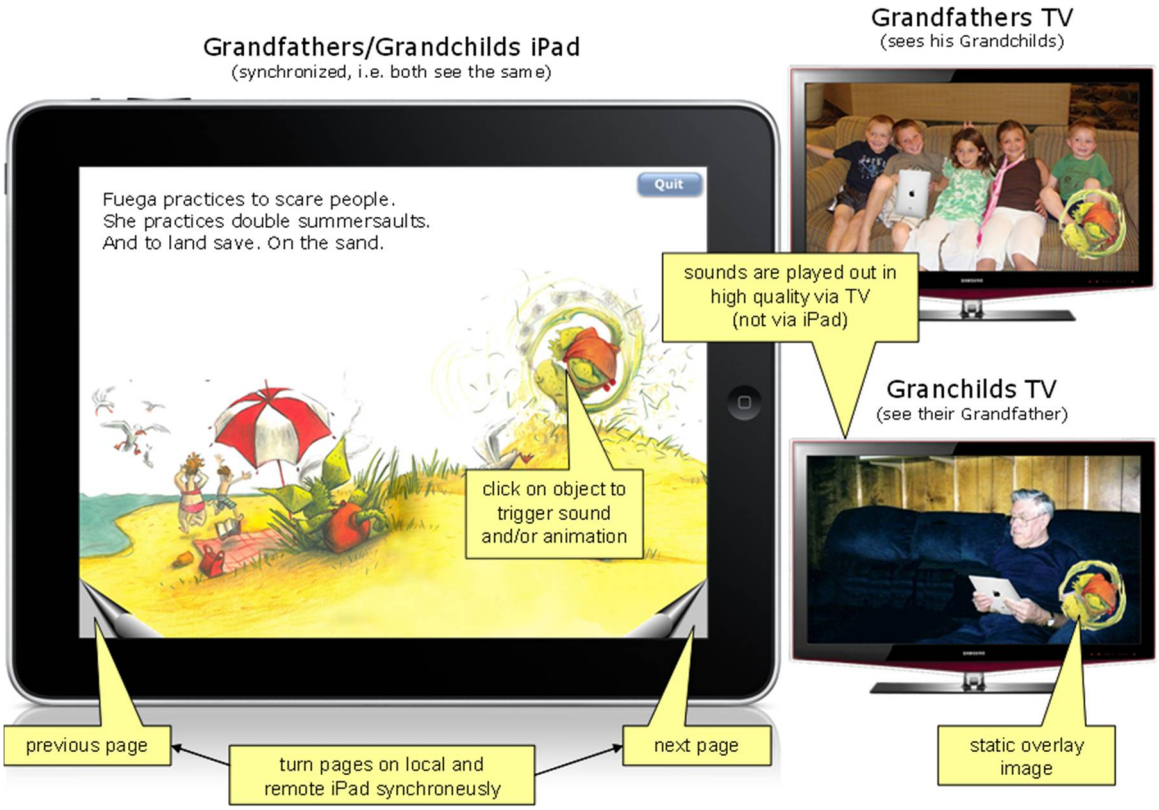


Figure 6 Illustration of the *Storytelling* demonstrator

3.1 *Storytelling*: the value of this demonstrator

Storytelling enables the project to conduct longitudinal in-situ trials that include children and older people (particularly grandparents) using synchronous interactions using good audio and high definition video. This is highlighted in Figure 1.

One of the underlying reasons for choosing the *Storytelling* demonstrator was that it could be enabled using relatively simple technology. This led to inherent stability and allowed the project to conduct long term in-situ user trials in people's homes and to thence learn about the higher level objectives of the project, namely "*making communication and engagement easier among families and friends separated in space and time*". Long term trials offer different insights from those that can be gleaned in a lab environment.

The *Storytelling* application uses a single HD camera and a stereo microphone. It connects two homes at a time, and uses an application that has relaxed real-time constraints. It embodies the fundamental features of a TA2 system: high quality communication and a shared activity.

Story telling as an activity between parents and children is particularly suited for in-situ trials because it can keep the interest over a longer period of time. Reading one chapter each evening can cause regular interactions over several weeks and therefore give time to improve a relationship. Furthermore, generating content for the *Storytelling* app on the iPad is relatively easy – at least compared to e.g. a mini-game in the Family-Game demonstrator that is gesture-controlled through audio and video analysis. Though such mini-games are very entertaining, it is time consuming and challenging to make them work reliably in an uncontrolled environment.

3.2 *Storytelling*: concept development

As illustrated in Figure 7, only 12 month were available from the first sketches on paper until the units had to be deployed in homes. Despite the ambitious goal, the milestones were achieved with a delay of only 4-8 weeks (Figure 7 shows the original project plan w/o the delays). Twelve *Storytelling* systems have been developed and these are being deployed either in house pairs (or triples) over 4-6 week evaluation periods. This will provide excellent rich qualitative data that will be supported by quantitative usage data from system logs.

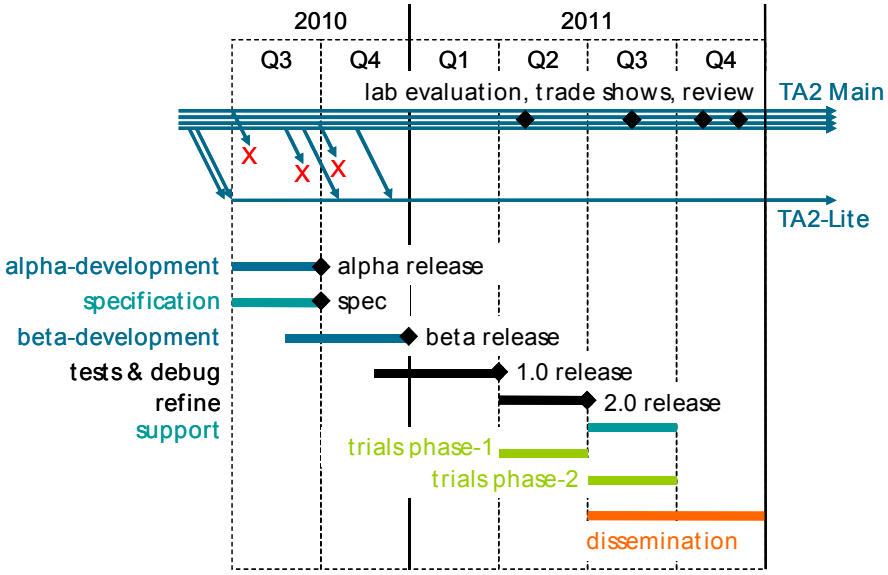


Figure 7 Original project plan of TA2-*Storytelling*

3.3 *Storytelling*: Evaluation results

During the first phase of user trials, systems were installed in a meeting place for elderly people and in the home of initial test families, see Figure 8. A technical staff member was present during these tests to make sure that the users could operate the system. Initially, the TA2-*Storytelling* units were set up in two rooms at the same location. Later on, two meeting places in the same city were connected via broadband Intranet. These early user trials were designed as a participatory design cycle. Users were welcomed to play with the system, to make comments and to give feedback. This pre-test revealed several bugs and usability issues which were (where possible) addressed during the development of the Release 2.0 of *Storytelling*.

Figure 8 Test installation in the home of a Swedish family during Phase-1 of the trials

Release 2.0 of TA2-*Storytelling* was delivered at the beginning of August 2011 and is being used for the second phase of user trials, in which the system is put into homes of families without the technical staff being present. Technical support via phone and at homes is provided if required. User trials in phase-2 are more like a formal experiment. By recording usage data and performing questionnaires/interviews results for qualitative and quantitative evaluation are collected. However, because the sample size will be small (about 4-8 pairs of families), the collected data may not be sufficient to provide statistically significant results. Hence, the focus will be on qualitative analysis in form of semi-structured interviews.

The main quantitative data captured is the usage frequency and duration. This data can be detailed into the usage of the videoconferencing system and the *Storytelling* application. For the former, the start and stop times of the videoconferencing sessions, as well as quality metrics, are recoded. For the latter, the download times of each book page as well as the usage of interactive elements (sounds, overlay, and animations) are captured. Based on this data we hope to evaluate how often the shared application is used, together with the videoconference, or if videoconferencing is rather used on its own most of the time. It can also be evaluated whether the books are read completely (indicating interest) or aborted in the middle (indicating disinterest), or what role interactivity and multimedia elements plays. Figure 9 show data of how a book is read during a test session. As can be seen, the chapter was read almost linearly until the end (page 22) after which it was turned back to page 15 to explore a detail of particular interest. The book chapter kept the users interested for about 7 minutes.

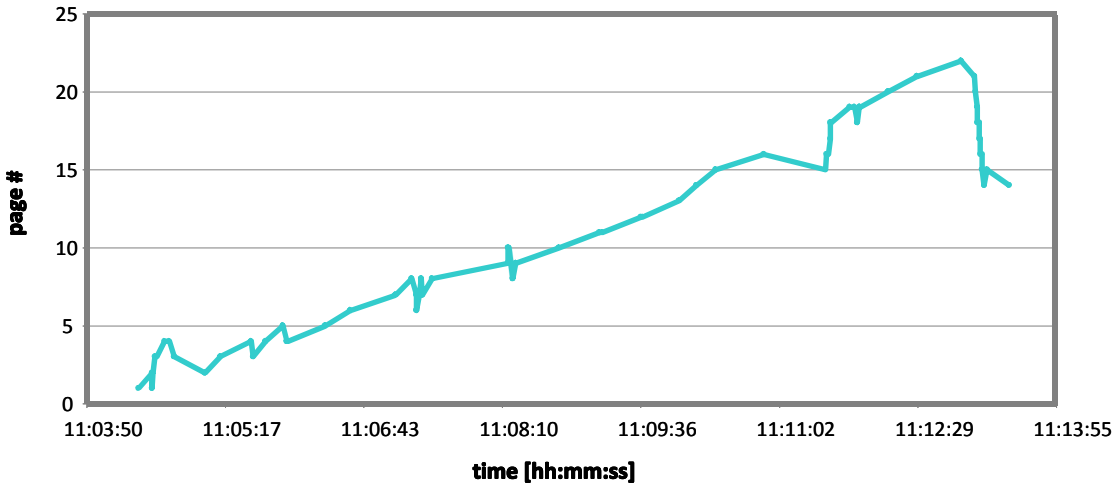


Figure 9 Example of quantitative data as will be recorded in Phase-2

One anecdotal but encouraging result has been that, without exception, all those participating in the trial have asked whether they can keep the TA2 set up. This suggests that something about the system was attractive but it is not clear whether this was just the iPad or whether it was the communication opportunities the combined system afforded.

3.4 *Storytelling:* What next?

The user trials in Phase-2 are still taking place and data is being collected. We have not obtained enough qualitative results from semi-structured interviews, nor have we obtained enough log files to start quantitative evaluation. We expect to collect data up until the end of 2011, and that evaluations can be completed by the end of January. From the analysis we hope to learn how TA2-technology is used in homes and if it actually can help to enhance relationships.

Figure 10 HD-Webcam and Stereo-Microphones integrated in a custom aluminium housing

There is a lot of tacit knowledge that is already gained through user feedback and trials conducted by Interactive Institute. This knowledge is difficult to capture and quantify but an example may help to understand the nature of the issues: The videoconferencing system is using a *TA2-Sensor*, which combines an HD-webcam and stereo microphones in an aluminum housing, see Figure 10. It was carefully designed by Fraunhofer IIS to assure correct placement of the microphones according to the ORTF specification (17 cm distance, 110 degree angle). However, when used in homes it turned out that the camera often needs to be placed above the TV, while the microphones should be closer to the user. To make things worse from an audio-expert point of view, the stereo-pair was taken apart as well and used to better record people sitting in different places of the living room. Though this completely destroys any stereo-image, one has to accept that it may actually better serve the need of users in the given situation. Hence, the complete housing-design can be put into question. Fraunhofer IIS can learn from this observation in several ways and e.g. support multiple-mono microphones in the Audio Communication Engine (ACE) in addition to a stereo-pair (and not go into the business of building sensor housings).

4 *MyVideos*

Description

The *MyVideos* demonstrator explores asynchronous social sharing and *Storytelling*. Using a school concert as the guiding theme, the demonstrator allows short video clips filmed by audience members to be compiled automatically into personalized stories that can be shared. We believe that using all the available clips, together with intelligent editing techniques, will improve the aesthetic and narrative quality of the videos created and that personalized presentation will make a more effective and stimulating catalyst in the building of relationships.

Technology

The demonstrator uses a number of the key capabilities in the TA2 architecture. Content Analysis is used to automatically match all the video clips against a common timeline and to help annotate them, for example by identifying people in the clip and the shot type.. Several tools are provided to allow users to explore the annotated video using both guided and fully-automated techniques – such as a recommender system to suggest related clips, and a rule-based narrative engine which automatically creates story playlists. Components from the Visual Composition Engine are used to dynamically assemble these video clips, allowing real-time interaction with the story.

Social Science

Storytelling is a fundamental part of the way we communicate; and video is a central way we consume stories as entertainment. Yet people struggle to translate the hours of video they capture into stories which can prompt communication. *MyVideos* is testing the assumption that technology can help people to create videos that are better, in terms of their narrative and aesthetic quality, and that these can lead to richer, more frequent interactions between people that know each other well.

Stories we tell based on our experiences are critical to recalling and reinforcing the memory of shared experiences that is so important to the nurturing of social groups.

Evaluation

We are evaluating both the concept demonstrator itself and the assertion that we can use technology to help nurture social relationships. The demonstrator is being evaluated through user tests with observation and qualitative feedback driving subsequent design decisions. The impact that the technology may have on social relationships is evaluated through a series of structured questionnaires comparing perceptions of the *MyVideos* demonstrator with existing resources. Three major evaluations have taken place over two and a half years with the software undergoing significant developments following each evaluation.

Learning

Users of the early *MyVideos* system were (generally) pleased with the personalised presentations that they produced and believed that such systems would encourage them to capture and share more video. This suggests that systems like My Videos will help nurture relationships between people that know each other well.

A number of features in the following W3C standards are significantly influenced by this work: SMIL Text (for animated captions), SMIL State (for incremental, on-demand media documents), and HTML Timesheets within HTML5 (for the development of synchronization technologies within common media browsers).

The *MyVideos* demonstrator examines togetherness in the context of the asynchronous social sharing of personalized media. Where most of the other TA2 demonstrators study communication and interaction within the context of a synchronous experience (where all of the participants are active at the same time), *MyVideos* looks at situations in which participants are separated in both space and time. This time lag allows us to further study personalized sharing: a given social experience can be tailored to the individual needs of participants based on an 'interest' profile, as well as on the profile of the hardware available or the social setting in which the sharing takes place.

The guiding theme of the *MyVideos* demonstrator has been that of a high school concert. The concert is an event by non-professional musicians, yielding presentations that often only a mother could love. The aim of social sharing within *MyVideos* is to help family members and other interested parties get the feeling of the event, in general, and to highlight the role of an individual child performer, in particular. This makes *MyVideos* fundamentally different from a common concert video mash-up. The main focus is the role of an individual performer in the event, rather than the event itself.

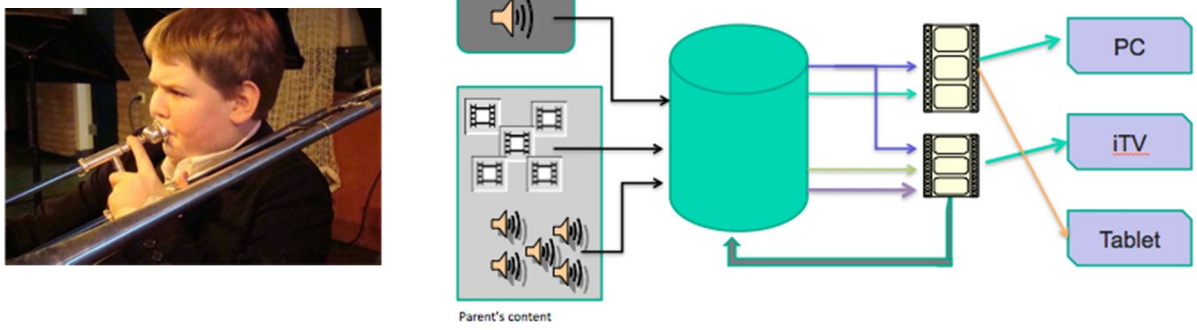


Figure 11 Representation of the *MyVideos* demonstrators illustrating how images from a concert are ingested into a shared repository of files in order to create personalised presentations that can be viewed on a PC an Internet TV or a Tablet

There are several technical challenges that *MyVideos* has studied. First, we have considered content analysis during the ingestion phase of audio/video processing. This has resulted in innovative tools and algorithms for analyzing meta-information within video content, including shot analysis, person identification and instrument identification. It has also resulted in efficient algorithms for content alignment based on matching the audio data in multiple fragments. The second set of technical challenges has included the development of a concert narrative for personalized delivery based on high-level and low-level characteristics of the concert and the musicians. The development of the resulting story has been developed in two threads: an automatic video general path, in which users are required to supply only limited control inputs prior to video generation, and a more manual and user-directed interface, which gives the ability for extensive user customization of the presentation's structure. Finally, we have studied the targeted delivery of content to various classes of user groups: the musician him/herself, his/her parents, an uncle living apart from the family and a remote grandmother.

The overall goal of the demonstrator is to determine if customized presentations provide the remote participants with an increased feeling of social togetherness. Given the two threads, we have compared the amount of work required to produce satisfying, customizable videos. We have conducted a series of trials with various generations of the demonstrator to gauge effectiveness, and have worked with a

constant group of parents throughout the development of the demonstrator in a participatory manner to define user-inspired interfaces and control structure.

4.1 *MyVideos:* the value of this demonstrator.

MyVideos enables the project to probe the value of a composition engine, able to control the temporal and spatial rendering of videos using the SMIL composition language for groups of people who already enjoy strong social ties through asynchronous interactions. This is highlighted in Figure 1.

Within the TA2 family of concept demonstrators, *MyVideos* is an application-level evaluation platform that helps us gain an in-depth understanding of interpersonal interaction within a wide range of social interaction. From a technological perspective, *MyVideos* makes use of the TA2 VCE and content analysis infrastructure, and integrates portions of the presence and orchestration functionality used in other TA2 demonstrators. The demonstrator has been designed for lab-based evaluation and for in-situ deployment across multiple endpoints, both for individual users and for social groups. The trials have been directed toward teenagers and their parents and grandparents. *MyVideos* was defined to allow TA2 to study asynchronous togetherness within the suite of project demonstrators. Within this general domain, studying sharing in the context of a community example would provide us valuable insights into how people would interact with time-dependent content in the context of long-term use and adaptation.

Media capture is a ubiquitous activity. Shared viewing of that captured content typically is not. In spite of the dramatic impact of user contributed content sites (such as YouTube and various national video sites), the amount of ultra-personal content being shared with family and friends (to say nothing of wide anonymous audiences) is minimal. A conservative estimate of media use indicates that average owners of smartphones and portable cameras capture hours of videos yearly, but that only minutes (or seconds) of content are being shared. The question we have is: why is this?

Modern video cameras allow a user to exercise a reasonable level of creative control during video capture. Shots are created using cameras with high resolution, often resulting in HD video output. Lenses are sensitive enough to allow capture to take place in dimly lit venues, yet small enough to fit in a shirt pocket. In general, audio quality is high, and recording skew is minimal. Finally, most cameras allow a dazzling array of metadata to be captured, including everything from camera details, locations and faces to be identified during capture.

In spite of these advances, content is not becoming more accessible: where face/person recognition in images is becoming commonplace, the successful high-level analysis of video content is still in its infancy. While many cameras support low-light capture, the result is that indoor video often is poorly lit and densely shot. This makes video analysis difficult, which in turn makes indexing, finding and reusing captured content by end-users frustrating.

The problems of (re)using personal content have several other dimensions as well. At a high level of abstraction, a fundamental problem is that third-party users expect a higher level of production content (in terms of shot selection, story pacing and logical narrative) than most amateur authors have available to them. At a lower level of abstraction, searching through captured content is simply too laborious to provide a productive user experience. To put it another way, as the digital family archive gets larger, the more difficult it becomes to find something interesting.

Within TA2, the *MyVideos* demonstrator has allowed us to gain a better understanding of how personalized media sharing can increase the social connectivity among people, even when they are not active at the same time. The *MyVideos* demonstrator is not a shrink-wrapped application, but a set of technology components that are used to determine selective aspects of shared asynchronous communication.

4.2 *MyVideos*: Concept development

The *MyVideos* demonstrator was developed at the start of the TA2. In the initial phases, we invested considerable time in working with parents and remote family members to understand the application-related needs of individual users. We wanted to find a combination of demonstrator complexity, evaluation structures and users' needs that would help us obtain useful results in the project.

The results of the demonstrator were developed as part of a long field study and trial, from late 2008 until November 2011, in which groups of families in the Netherlands and the United Kingdom collaborated in preliminary interviews, video recording and social sharing of content within a relatively closed community. Our long-term association with these families has provided us invaluable insights on how so-called Web 2.0 technologies fail to meet the demands for interpersonal communication between friends and families. It has also provided a deep understanding of current practices around home videos, and how such practices can be matched to existing social theories.

In order to better understand the problem space, we involved a representative group of users at the beginning of the evaluation process. The first evaluation, in 2008, consisted of interviews with sixteen families across four countries (UK, Sweden, Netherlands, and Germany). The second evaluation, in-depth focus groups – with three parents each – was run in the summer of 2009 in the UK and in December 2009 in the Netherlands. Since our requirements for socially-aware video editing and sharing systems are partially derived from the results, we will summarize here key findings about current practices and challenges about media sharing.

In line with the models of Durkheim (Durkheim, 1971) reliving memories and sharing experiences help bring people together. This creation of memories of shared experiences is central to the *MyVideos* demonstrator. Parents e-mailing pictures of the kids playing football to the grandparents, or sharing holiday pictures via Picasa, or on disk, or using Facebook, enable friends and families to stay in touch with each other's lives. Nevertheless, the interviewed people said that if they shared media, they would do so via communication methods they perceived as private and then only to trusted contacts. There was a general reticence from the parents towards existing social networking sites. From the focus groups it became clear that current video sharing models do not fit the needs of family and friends. Much richer systems are needed and will become an essential part of life for family relationships.

The focus group comments helped form the foundation for demonstration used in the *MyVideos* trials. These trials consisted of initial media capture and analysis of events in England and the Netherlands, followed by two in-school trials at a Dutch high school. A final trial event will be held in the UK in November, 2011.

In the development of *MyVideos*, we developed two parallel story construction models: *automatic authoring*, in which the TA2 system constructs a video dynamically, and user-directed *manual authoring*, in which a power user is given control during story development.

4.3 *MyVideos:* Evaluation results

The first set of *MyVideos* evaluations were developed using user-directed manual authoring of story content. As is illustrated below, (Figure 12) such editing is particularly useful when constructing highly-personalized summaries at the performer level. An evaluation of more automated authoring aspects is scheduled for late 2011.

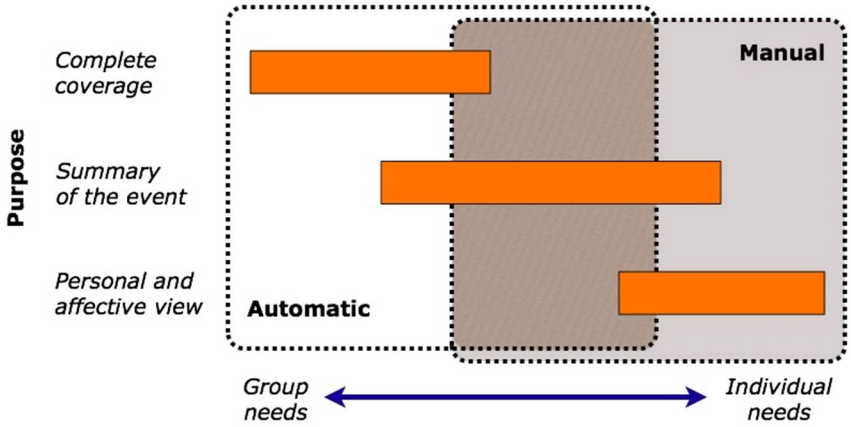


Figure 12 A figurative representation illustrating how the presentations of a shared event like a school concert are different for individuals and groups

The high-level workflow of initial *MyVideos* demonstrator is sketched in the illustration below (Figure 13). The input material includes the video clips that the parents agreed to upload, together with a master track recorded by the school. All the video clips are stored in a shared video repository that also serves as a media clip browser in which parents, students, and authorized family members can view (and selectively annotate) the videos. Privacy, and a protected scope for sharing, is a key component of the system. Each media item is automatically associated with the person by whom it was uploaded, and there are mechanisms for participants to restrict sharing of certain clips. Participants can use their credentials for navigating the repository – those parts allowed to them – and for creating and sharing different stories intended for different people.

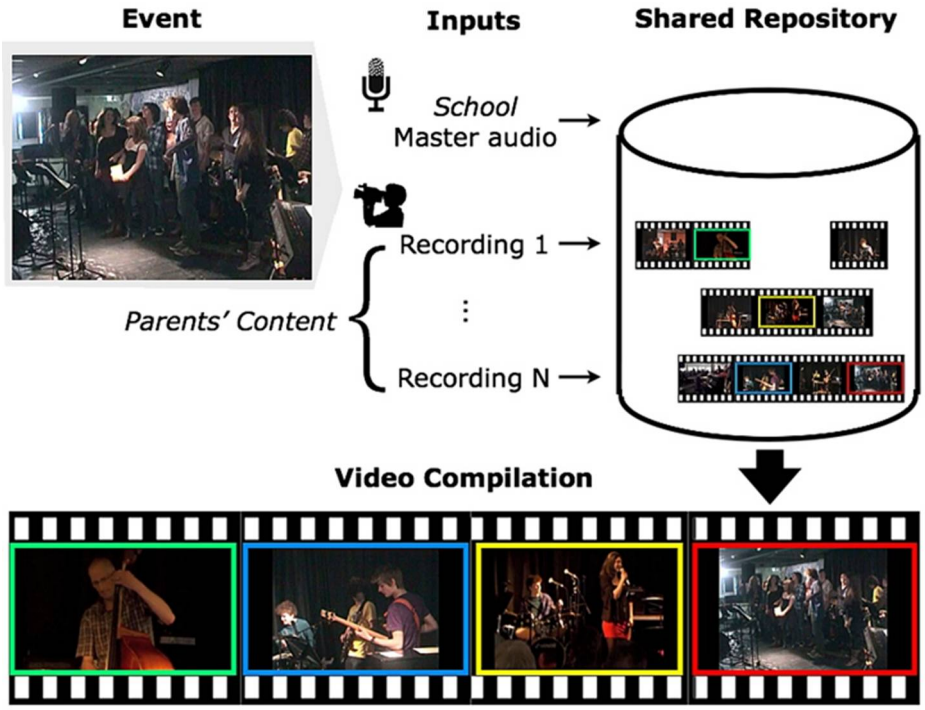


Figure 13 Illustrating how clips from multiple recordings are ingested into a shared repository and then played out as a compilation against a timeline.

Users are not told what they should capture, they can record what they wish using their own camera equipment. The goal is to recreate a realistic situation, in which friends and families are recording at a school concert. This flexibility comes at a cost, however, since most existing solutions that work well in analysing input datasets are not that useful for our use case. Handling user-generated content is challenging, since it is recorded using a variety of devices (e.g., mobile phones), the quality and lightning are not optimal, and the length of the clips is not standard.

Apart from allowing fine-tuning of assembled videos, our system enables users to perform enrichments. *MyVideos* provides mechanisms for including personal audio, video, and textual commentaries. For example, these could be recordings of oneself (the author of the production) or subtitles aligned with the video clips commenting the event for others. Users can also record an introductory audio or video, leading to more personalized stories.

Our application addresses reciprocity by enabling life-long editing and enriching of compiled videos. Videos created with our tool can be manually improved and enriched using other assets from the repository, adding personal video and audio recordings, and including subtitles or textual comments.

4.3.1 Video Recordings

The first set of experiments were based on recordings of three different concerts: a school rehearsal in Woodbridge in the UK, a jazz concert by an Amsterdam local band called the Jazz Warriors, and a school concert at the St. Ignatius Gymnasium in Amsterdam. The intention behind these concert recordings was twofold: to better understand the problem space and to gather real life datasets that could be used for testing the functionality.

In the first concert, Woodbridge, a total of five cameras were used to capture the rehearsals. The master camera was placed in a fixed location, front and centre to the rehearsal, set to capture the entire scene (a 'wide' shot) with no camera movement and an external stereo microphone in a static location

near to the rehearsal performance. This experiment was very useful for testing the automatic processes provided by our system: the temporal alignment algorithm, the semantic video annotation suite, the automatic authoring tool, and the recommender system for visualizing multi-camera recordings.

Then, at the end of November 2009 a concert of the Jazz Warriors in Amsterdam was captured as part of an asset collection process. The goal of the capture session was to gain experience with an end-user setup that would be similar to that expected for the Amsterdam school trial. The concert took place on 27 November 2009 at the *Kompaszaal*, a public restaurant and performance location in Amsterdam. The Jazz Warriors is a traditional big band with approximately 20 members. In total eight (8) cameras were used to capture the concert, where two cameras were considered as 'masters' and were placed at fixed locations at stage left and stage right. In total, we collected more than 200 video clips and approximately 80 images. The longest video clip was 50min, the shortest 5s.

The first two concerts were primarily experimental, providing us enough material for fine-tuning the automatic and manual processes. On 16 April 2010 the concert from the Big Band (St. Ignatius Gymnasium) was recorded. In this case parents took part in the recordings and provided the research team with all the material. In total around 210 media objects were collected for a concert lasting about 1h and 35 minutes. Twelve (12) cameras were used; two of them used as the master cameras. These recordings were used for evaluating the prototype.

4.3.2 Prototype Evaluation

In cooperation with the local high school, a core evaluation group of seven people were recruited among relatives and friends of the young people that performed in the third concert. The rationale used for recruitment was to gather as many roles as possible, in order to better understand the social needs of our potential users. The participants were three high school students, a social scientist, a software engineer, an art designer and a visual artist, resulting in a variety of perspectives that may influence the video capturing, editing and sharing behaviours. All the participants were Dutch. The average age of the participants was 37.1 years (SD = 20.6 years); 3 participants (42.8%) were female. Among our participants, 3 had children (ranging in age from 14 to 17 years). All participants were currently living in the Netherlands, but an uncle of a performer living in the US was recruited to serve as an external participant (the only one that was not present in the concert recordings). The prototype evaluation was conducted over an eight week span between July and September of 2010.

The review group was kept small so that we could establish directed and long-term relationships. The qualitative nature of our interactions provided us with a deep understanding of the ways in which people currently share experiences to foster strong ties. The participants represent a realistic sample for the intended use case: all the parents have kids going to the same high school, all of them tend to record their kids, and some of them have some experience with multimedia editor tools. Moreover, the parents were involved in previous focus groups dating from December 2009 and they recorded their kids playing in the Big Band concert (April 2010). The goals of the study make it impossible to do crowd-source testing, since users of the system should be people that care about the content of the videos.

We used multiple methods for data collection, including face-to-face interviews, questionnaires, and interaction with the system. At the simplest level, users only had to select the subject matter (people, songs, instruments) and two parameters (style, and duration). Then, by pressing a "GO" button, a story was assembled.

In addition to automatic creation, users could also browse the media repository to view the collection of objects contributed by parents. The image below (Figure 14) shows one interface for temporal browsing. In the figure, the video clips in which a performer connected to the user appears are highlighted in orange. Note that browsing could focus on events, cameras, people or instruments in the database.

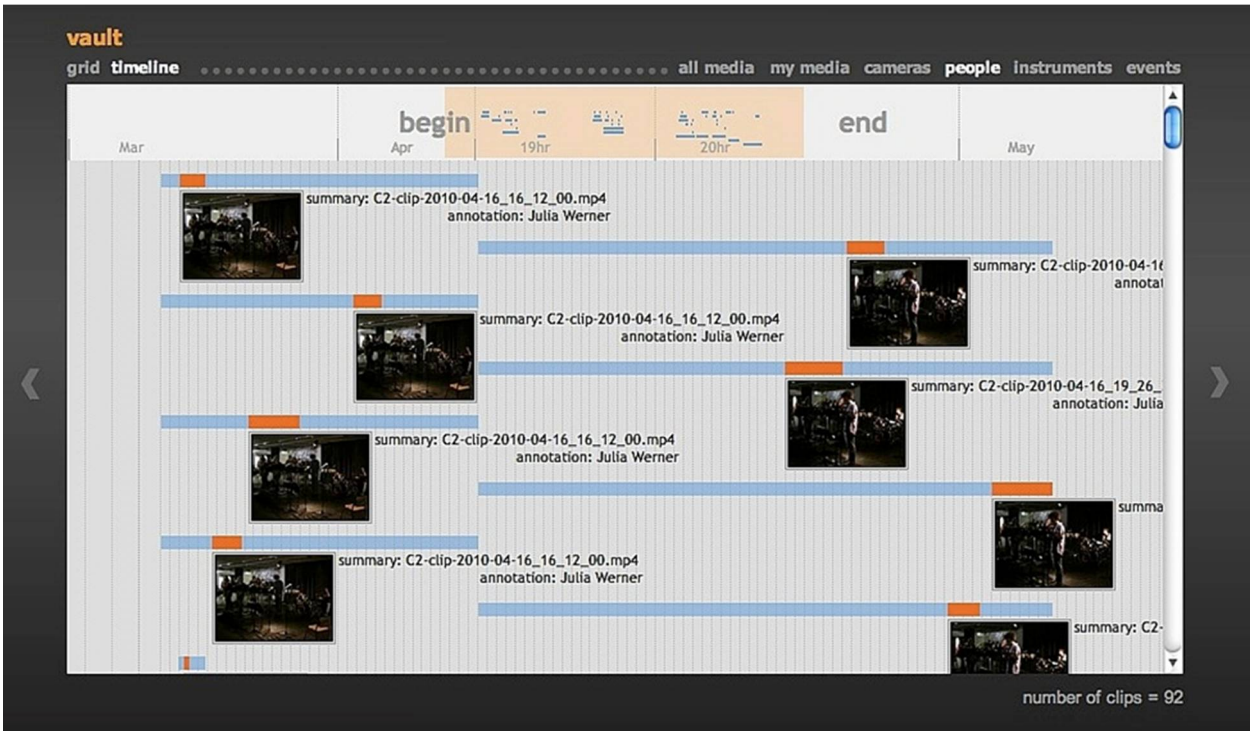


Figure 14: A normalized timeline view of sparse fragments contributed by parents, after audio alignment by *MyVideos*

The evaluation experiment consisted of 2 sessions. The initial session was used to collect background information about video recording habits, their intentions behind these habits and their social relations around media. We also used this session as an opportunity to understand how the participants conceptualize the concert. The second (in-depth) session was dedicated to capture video editing practices and media sharing routines of the participants based on their interactions with the system.

4.3.3 London Student Study

Another smaller and more incidental evaluation was performed for validating and contrasting the results obtained using the prototype. In this case, we asked a broader group of British students about their social media habits. We then introduced the system to them, demonstrating both media browsing and automatic compilation. A total of 20 young people completed the questionnaires. The students were from London high schools and were between 14 and 18 years old.

4.3.4 Results

Most of the results that were obtained using the demonstrator were qualitative and were reported in project deliverable D7.14. Some of the results were compiled via participant questionnaires. The illustration below (Figure 15) shows the results of some of the questionnaires filled by the participants in Amsterdam, highlighting their answers to the questions related to utility and usefulness, including comparisons to other existing solutions.

Questions		Questionaire Results
Q01. Did you like *MyVideos*?		
Q02. Does *MyVideos* help you recalling memories of social events?		
Q03. Would you like to see what other parents and friends have recorded in the same event?		
Q04. Is *MyVideos* better than traditional systems to browse videos recorded by other people?		
Q05. Is *MyVideos* better than traditional systems to find people you care about?		
Q06. Would you add/correct more metadata by using *MyVideos*?		
Q07. Is the material you capture enough to create video productions?		
Q08. Would you create more video stories with *MyVideos* (compared to existing tools)?		
Q09. Would you create videos stories faster using *MyVideos* (compared to existing tools)?		
Q10. Would you create video stories more easily using *MyVideos* (compared to existing tools)?		
Q11. Would you fine-tune automatic generated video stories by using manual tools?		
Q12. Would you personalize videos by adding yourself to the story? ("capture me")		
Q13. Which approach did you like the most?	Automatic	
	Manual	
Q14. Would you share more videos with *MyVideos*?		
Q15. Is *MyVideos* safer than other video sharing systems?		
Q16. Would you pay for *MyVideos*?		

Key: (Yes) (No)

Figure 15 Questionnaire results from the Amsterdam trials, highlighting questions related to utility and usefulness

We can summarize these responses as follows. Regarding common practices, all participants reported that they record videos in social events (e.g., family gatherings and vacation trips). However, most of them said that they barely looked at the recorded material afterwards. For most of them, video editing was time consuming and overly complicated. Nevertheless, some were familiar with video editing tools. When asked if and how they share their videos, our participants repeatedly said that in general they do not post personal videos on the Web. While the youngest participants argued their personal videos were not interesting enough to share on social networking services, our older respondents cited privacy concerns as the main reason not to share personal videos on the Web.

The responses from the London high school questionnaires are shown in the graphs in Figure 16. In general, the participants in London came from a generation that has grown up watching and sharing digital videos. Regarding *MyVideos*, the survey participants graded it similarly to the evaluators of the prototype, indicating that our results may be generalised.

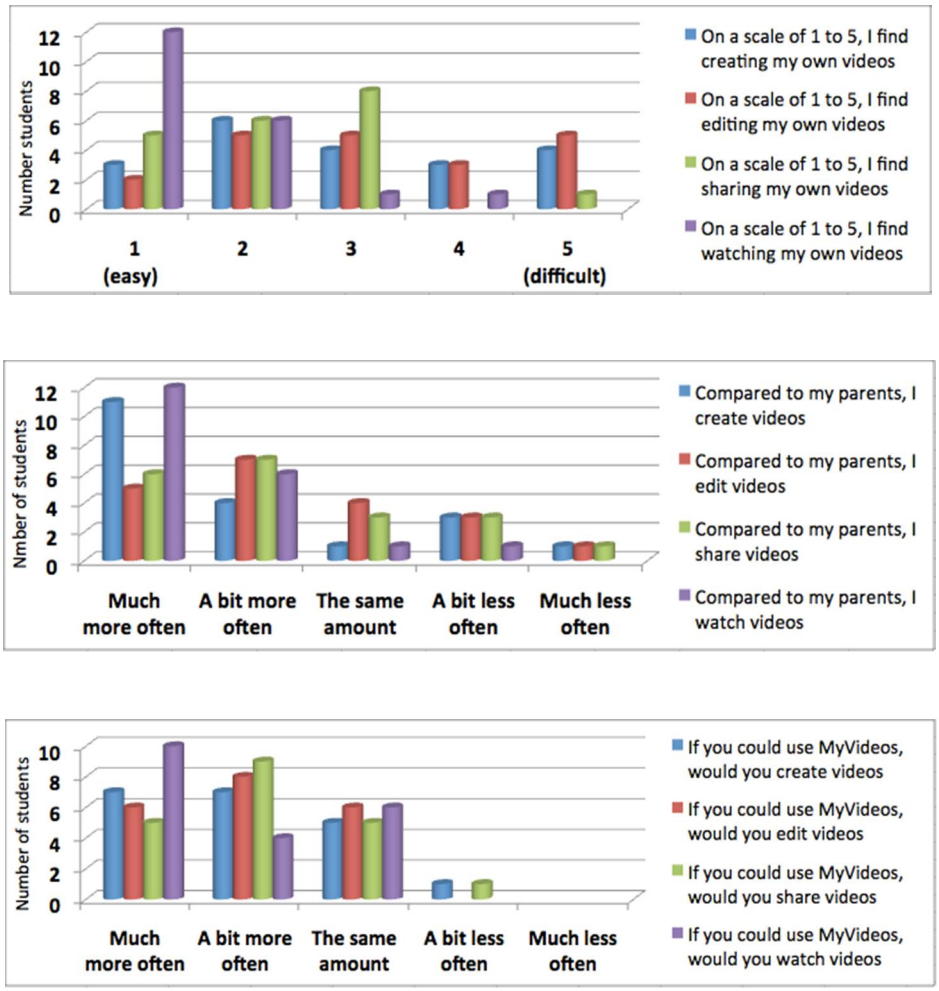


Figure 16 Results from questionnaires completed by London school children revealing some of their reported use of video and their perceptions about the *MyVideos* demonstrator

4.4 *MyVideos*: What next?

A final trial of the *MyVideos* demonstrator is scheduled for late November in England. The goal of this trial is to evaluate automated story construction and distribution. It will also evaluate new methods for constructing shared video experiences directly from the browser interface. These results will provide comparison data for the baseline results generated during the Dutch tests.

MyVideos has proven to be a successful vehicle for testing a variety of user and TA2 assumptions on the way in which social sharing can be enhanced by using personalized asynchronous presentations. While we have focused the development of the demonstrator on building a system that was evaluation-ready, we feel that the concept as a whole can serve as the basis for a potential exploitation path. Servicing the home market (as was done for the concert theme) may be unrealistic, but when combined with a vertical market – such as is being done in Music Tuition – we feel that *MyVideos* has many reusable components.

Experiences gained during *MyVideos* led to the integration of a number of features in the following W3C standards: SMILText (for animated captions), SMILState (for the development of incremental, on-demand media documents), and HTML Timesheets within HTML5 (for the development of synchronization technologies within common media browsers).

5 The *Connected Lobby*

Description

The *Connected Lobby* is an application by which remote parties can join an HD video communication session from their TV screen and also the means by which they can discover and initiate the different sorts of experience that the TA2 system enables. The key research we are undertaking is to explore how such presence information for groups can be managed and displayed on shared devices such as TV screens and picture frames.

Technology

The presence and session control aspects of the *Connected Lobby* uses an array of largely proven technologies based on protocols like XMPP and Jingle. A key contribution from TA2 is to adapt the application so it is suitable for use on a TV screen. This has involved developments in the way presence information is presented for groups in a social space, rather than for individuals on personal screens as is the norm. In addition passive displays, such as simple coloured lamps, have been used to indicate presence and also to act as the visual equivalent of a telephone ring tone for a video call.

Social Science

The *Connected Lobby* increases a sense of social presence in two ways. In line with with the meaning adopted from computer mediated communication, it acts as the personal stamp indicating that an individual or group is available and willing to engage and connect with others in their online community. Because the *Connected Lobby* initiates HD video communication, it also affects social presence in the context of real time communication, where "the degree of social presence is equated to the degree of awareness of the other person in a communication interaction. HD video should provide greater awareness than a system with lower quality video or no video at all, certainly for visual communications such as facial expression, gesture and body language.

Evaluation

Design concepts for interfaces that represent the availability of groups of people addressable through a shared device (the TV) have been evaluated in lab-based user tests. The use of ambient lights to act as 'ring tones' for a video chat were assessed through in-situ user trials. Further lab studies with up to 50 users are planned to evaluate the refinement of the *Connected Lobby* interface developed for a Philips NetTV.

Learning

User studies with families highlighted numerous anecdotal concerns about privacy which have informed our design and led to the inclusion of user controls to help users balance their needs for both awareness and privacy.

In-situ studies with video communication also highlighted that people view a video session as more like 'going for a coffee' than making a phone call. The "event" perception of the video session is a useful finding which should be reflected by product designers in the way they build user interfaces and market services. Consumers were also positive about the ambient light interface as "a very friendly way of passing information that can be silently ignored, unlike a ringtone."

The *Connected Lobby* forms a critical part of the overall TA2 concept. The *Connected Lobby* is the means by which remote parties can join a communication session and can discover and initiate the different sorts of experience that the TA2 system enables. As the Lobby name suggests, the lobby is traversed on the way to the TA2 experiences.

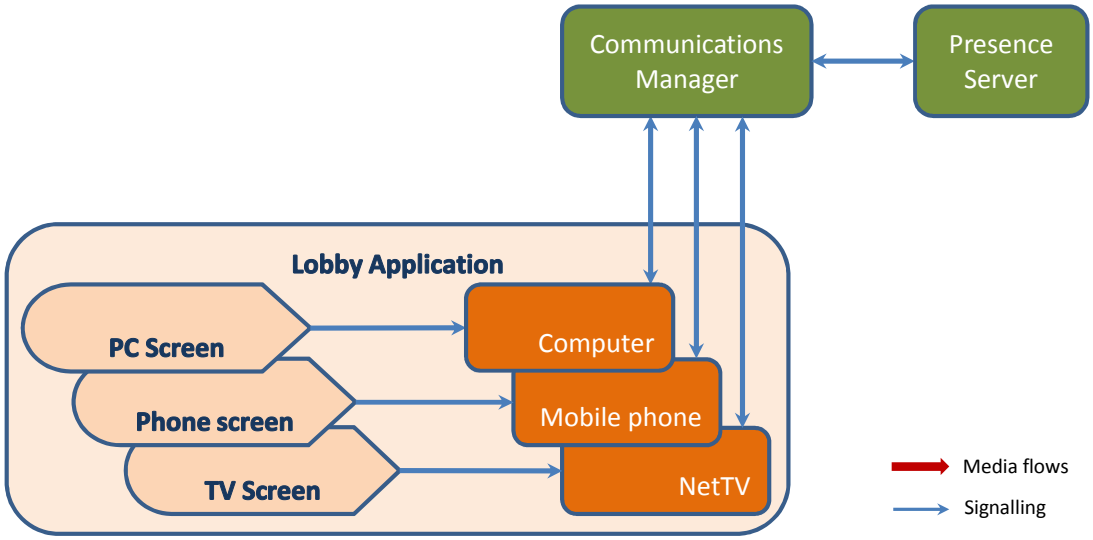


Figure 17 Conceptual representation of the *Connected Lobby*

The *Connected Lobby* is an application that (as depicted in Figure 17) has its output displayed either on a pre-existing screen or through a more ambient display such as a lamp. The key research we are undertaking is to explore how such presence information for groups can be managed and displayed on devices including shared devices like TV screens and picture frames. This is quite a different challenge to presence system related to communications clients like Microsoft Communicator or Skype which relate to individuals only.

The *Connected Lobby* should aid communication by increasing a sense of *social presence*. Indeed with its twin roles of managing presence and of initiating basic video chat, The *Connected Lobby* should affect *social presence* in two senses. Because the *Connected Lobby* manages and represents presence it should affect *social presence*, with the meaning adopted from computer mediated communication, i.e. as being the personal stamp that indicates that the individual is available and willing to engage and connect with other persons in their online community. Also because the *Connected Lobby* initiates HD video chat it should also affect *social presence* in the context of real time communication, where "the degree of *social presence* is equated to the degree of awareness of the other person in a communication interaction. HD video chat should provide greater awareness than a system with no video, certainly for visual communications such as facial expression, gesture and body language.

The *Connected Lobby* is used to display information relating to all users on the system, as well as provide access to the other applications in the TA2 project. It has the potential to display a whole range of information and media relating to specific users and also provides a platform for both synchronous and asynchronous communication between users. Also through sensing and displaying mood, availability, location, emotion etc., it is thought that a greater sense of awareness can be felt between family members.

Privacy and ethical issues regarding the display of information is obviously a major factor in a system such as this and have been investigated as well.

5.1 *The Connected Lobby*: the value of this demonstrator

One of the key challenges that the Lobby enables us to address is to understand how groups of users become aware of, and initiate interaction with, other groups of people. In the context of TA2 it is assumed that these groups area already known to each other but even so, intimate knowledge of the activity, presence or the communications resources available to the remote group are not easily discerned. This is a key function of the *Connected Lobby*.

Philips have led this concept and they are particularly keen to understand the role of the TV (and of lighting) in the representation of presence and in the initiation of group based experiences that are augmented by video, and delivered to a TV Screen.

As we are all aware Televisions are not simply devices for presenting your favorite shows and connecting your game and movie equipment. Products like the Philips Smart TV also support connections with wireless input devices (such as smart phones or tablets, which enabling one to share pictures and video) and also with the internet through which enhanced EPGs video stores and websites tailored for the television can all be accessed.

TA2 is keen to learn more about the role of the TV set as a communications device serving groups rather than just individuals. Recent research has shown that video chat via television can be highly valued in a family setting (Ames, 2010) (Ducheneaut, 2008.) Research in the context of social television has focused mostly on enhancing the viewing experience required for text chat and video calls (Geerts, 2006) (Harboe, 2008) (Huang, 2009). However, very few studies focus on the use of the television as a communication device by itself, being able to share various kinds of content.

Experiments using the *Connected Lobby* demonstrator suggest that using television as a social communication device has potential, not only for enhancing the television viewing experience, but also for simply communicating daily activities in a group setting (Demey, On usage and communication needs for the TA2 system, 2010).

5.2 *The Connected Lobby*: concept development

5.2.1 Initial thoughts

The *Connected Lobby* emerged from a demonstrator called “Enhancing Social Communications” that consisted of three scenarios, “Sharing Emotions”, “Chatting” and “Jigsaw Puzzle”. The first design-led ideas envisioned a large shared wall-sized display and focused on how different people could be represented and on what attributes one might choose (such as I am Happy) in augmenting the more prosaic information that presence systems provide today. This activity was very imaginative but, being based on wall size touch screen displays, the ideas could not be meaningfully evaluated in long term trials for cost and practicality reasons.

In late 2008, the TA2 project partners agreed that Enhancing Social Communications should be seen not as a stand-alone module but as an underlying ‘always on’ system capability that enabled communication and that triggered people to start other TA2 applications by providing representations of your family members along with pictures and text messages they have sent and of their current communication options they have open to them. The *Connected Lobby* can be regarded (see Figure 18) as an application but it should be noted that it is common to all the ‘other’ TA2 Demonstrators and that it depends heavily on control functions such as the communications manager and the presence server.

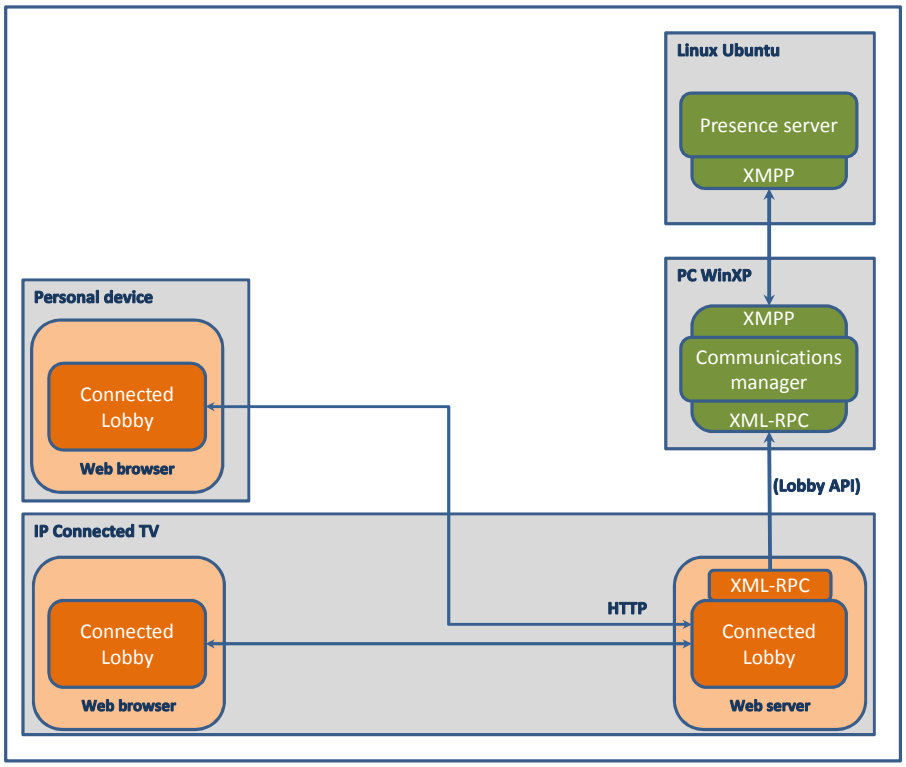


Figure 18 The *Connected Lobby*, as depicted within the overall TA2 Architecture Framework

We describe this Lobby as a "Connected" lobby because it may be accessed not only from the TV on which the TA2 activity will take place but also on any connected device – thus providing greater accessibility to information about the status and presence of your family circle.

The *Connected Lobby* is innovative as it sets out to support group-to-group interaction; design concepts have to provide ways of representing both groups (say 'the household') as well as individuals.

Functionally, the *Connected Lobby* allows synchronous and asynchronous information exchange between family members. It enables the transmission of pictures and of short messages from the different devices; its interface is available on many devices.

5.2.2 Flash demonstrator

The *Connected Lobby* was first implemented as an Adobe Flash application using Action Script 3 (AS3) running on a Windows XP 32-bit TA2-platform (see Figure 19).

The *Connected Lobby* was linked to the other components of the TA2 system via the Communication Manager (CM). It registered itself with, and could start other applications via Local Control (CM-LC).

Presence messages, as well as chat messages, were routed via the Session Controller (CM-SC). The CM-SC is responsible for session initiation and session management. The Lobby polled the CM-SC for updates on presence status or new messages; we chose this method as the Lobby acts as an XML-RPC client.

The Interaction Manager (CM-IM) component was a placeholder which represented how ancillary devices could be connected to the TA2 system if they were not an integral part of a particular application. Key presses and mouse events were directly captured from the Flash environment. It was

envisaged here that the CM-IM could receive messages from the Lobby, telling it to steer the LED controller of a light-augmented picture frame. This picture frame could be used as a TA2 notification device in case the display device (TV or PC) is turned off and unable to display presence information or incoming calls.

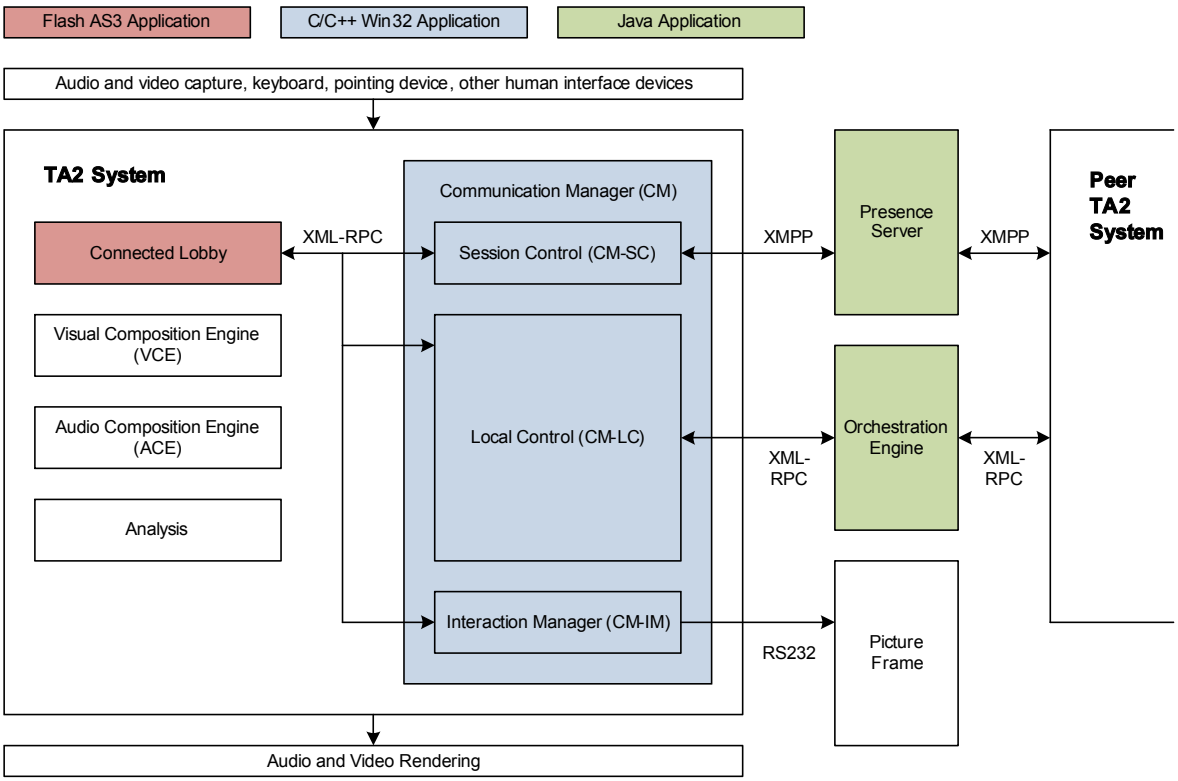


Figure 19 Schematic of the *Connected Lobby* showing the agreed communications interfaces between the *Connected Lobby* and the other components

The presence server is a LAMP server (Linux Apache MySQL PHP) running Ubuntu (a "human" Linux distribution) with Apache2 for HTTP serving, MySQL database server and PHP for scripting. OpenFire must been installed. OpenFire is an XMPP server that uses the MySQL database to store session data, user profile data, etc. Red5 is an OpenFire plugin that enables audio and video sessions via RTMP or SIP.

The Flash Lobby supported real-time audio and video communication via RTMP (a SIP-like protocol), the possibility to share messages and images, MUC (multi-user chat) as well as multi-user video conferencing, and the display of avatars.

The Flash Lobby was targeted at displays able to render Flash ActionScript 3 applications. In order to make the Lobby suitable for in-home testing, we wanted the user interface to serve multiple devices, e.g. iPads, mobile phones, wireless photo-frames, and Philips Net TV. Therefore we separated the Lobby's user interface from its back-end functionality. The front-end device can be basically any device running a web browser and it is fed by a *Connected Lobby* back-end, which is simply a web server.

5.2.3 *Connected Lobby* on a Smart TV

All Philips Smart TVs include Net TV. Net TV is based on global industry standards such as Open IPTV, which aims to define specifications for IPTV that will take the next generation of IPTV to the mass market. The forum is fully open to participation across the communications and entertainment industries. Furthermore, Net TV is based on the CE-HTML standard, which is a language tailored to show graphical user interfaces on consumer lifestyle devices.

Net TV has been available from 2010 and is available in all 7000, 8000, and 9000 series Philips TVs. It runs a CE-HTML browser on a chipset from NXP Semiconductors with no significant on-board storage. It is not possible to download and save web pages or install plug-ins. Figure 20 shows the *Connected Lobby* integrated with Net TV.

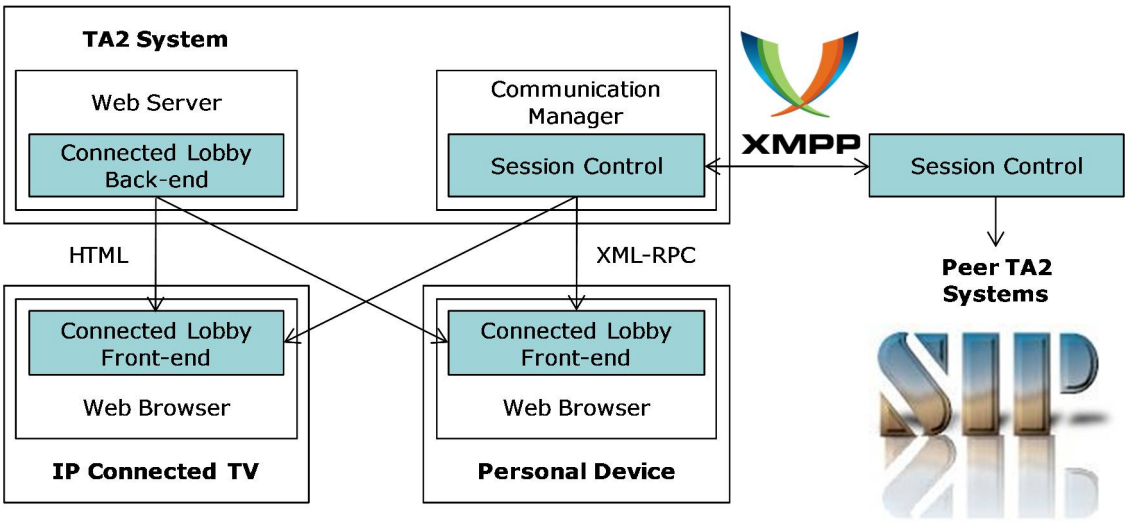


Figure 20. The *Connected Lobby* Architecture.

Figure 20 shows the new architecture of the *Connected Lobby*, where the user interface has been disconnected from Lobby back-end. The front-end communicates to the TA2 Communication Manager. The *Connected Lobby* UI can now be presented to multiple devices (Net TV, smart phone, tablet PC, etc.), so individuals can log in to the local TA2 system. It also allows Multiparty Multimedia Session Control based on Session Initiation Protocol (SIP).

5.2.4 Interface Design

With the integration of the *Connected Lobby* with Net TV its UI went through a complete re-design to make it easy to use, clear, and useable with a pointing device.

Figure 21. *Connected Lobby* interface design

Figure 21 illustrates the friendly, intuitive interface showing a notice board where family members usually share their messages. It still has the Net TV grid-like layout to enable navigation without pointing device, though it is not recommended.

The new interface supports group view of families. One can select a group, or individuals. The interaction possibilities are updated accordingly. Navigation is possible with a uWand pointing device, which also enables text input via full QWERTY keyboard.

5.3 *The Connected Lobby* : Evaluation results

The impact of migrating connected functionality, particularly to do with social media and social TV , from the set top box to the TV set is a key research question for television manufacturers like Philips and the work done in TA2 has provided valuable insights in a number of areas.

5.3.1 User Interfaces

To a certain extent, navigating through menus can be replaced by performing physical actions (Eichhorn, 2008). The idea is to use tangible user interfaces (TUIs): small physical objects showing the messages from family members. Each family member then possesses set of real physical objects, each representing one of his contacts. They could be equipped with a display, camera, microphone, light, and all kinds of sensors (Figure 22).

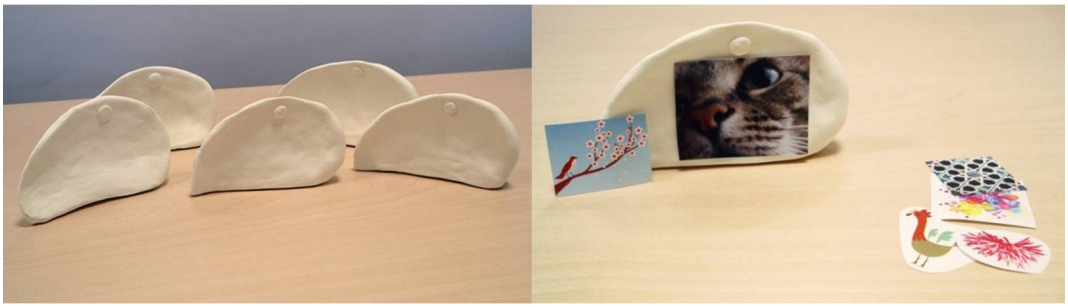


Figure 22. Prototypes of tangible user interfaces

Focus group discussions resulted in several advantages and disadvantages of using TUIs. If the objects and gestures are chosen carefully, these physical interactions have a high potential to increase usability, joy of use and create a new user experience. On the other hand, one has lying around multiple objects in the home, and the object's small size limits its functionality (e.g. keyboard).

When we use a mobile phone instead of a custom TUI we gain the main advantages and avoid the main disadvantages. Use of a mobile phone enables sharing experiences and being able to be connected anywhere and anytime and avoids "yet another device". As modern mobile phones have built-in cameras (often even two), microphones, light sensor, accelerometers and vibration motors, a lot of the gestures and interactions are still possible.

5.3.2 Balancing awareness and privacy

The main task of this research project was to explore the human factors relating to ambient intelligence in order to provide strategic design insights for the "*Connected Lobby*". The *Connected Lobby* was intended to enhance family togetherness whilst respecting concerns about privacy (Chen, 2009). The research focused especially on investigating the possibility of balancing togetherness and privacy through computer-mediated communication (CMC).

During the work it became increasingly clear that a family is a very sensitive group (Eichhorn, 2008) and that we had to be very careful with privacy issues. This led to a few key decisions such as

choosing NOT to visualize or provide any traces of conversation between certain family members. Visualising or providing traces /records of conversations is easy but to not do so matches what happens in real life and steers clear of privacy issues.

Privacy towards the system is also an issue. While collecting data we will use different sensory equipment (at least cameras and microphones for video calls). So the user has some kind of "intelligence" and data collector in his house that he has to be able to trust. Also a lot of the data will be transmitted and stored. Even if we know the system to be reliable and secure we cannot guarantee that users will trust it so sensitive decisions need to be made in choosing what data is collected, a typical guide should be to collect store and transmit the absolute minimum in order to minimise the scope for the development of user mistrust.

We do not expect a single solution to fit everyone's preferences and suggest that users should be given control so they can proactively choose what sorts of sensors are used.

In addition a generative session was organized to find out people's triggers for communication, their habits in choosing the medium and how to share personal information with family members and close friends

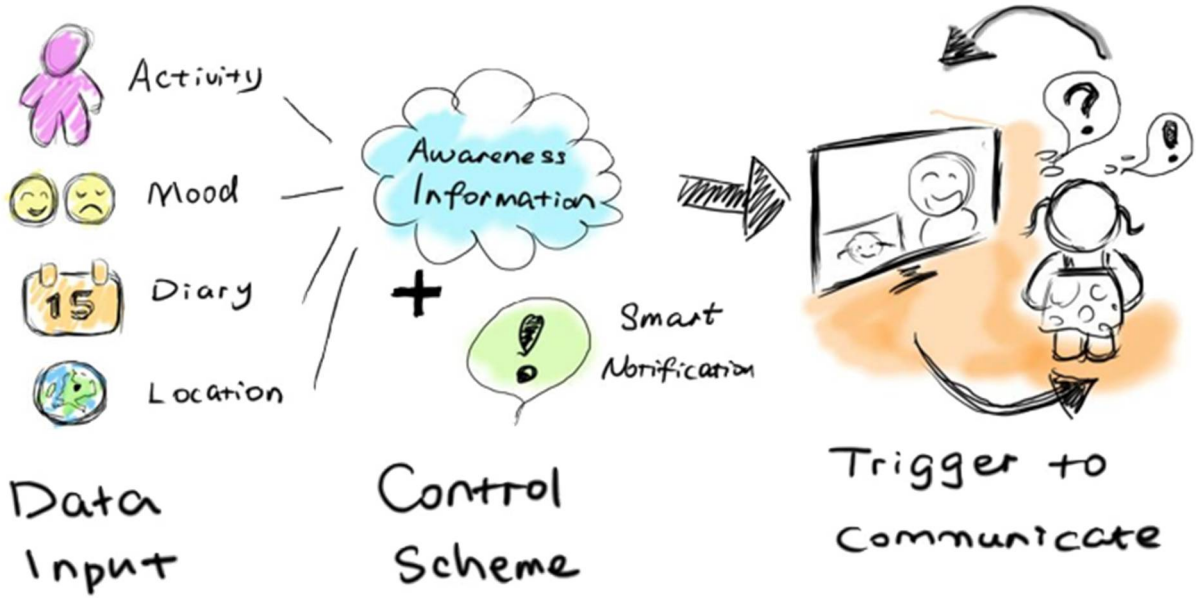


Figure 23. Smart awareness assistant

This led to a strategic design guideline on a control scheme, namely how to use ambient awareness systems to trigger people to communicate. The concept of a "smart awareness assistant" (Figure 23) was developed and a renewed "*Connected Lobby*" interface (Figure 24) was designed and tested. It was designed to address the following requirements:

- to enable people to talk to their family depending on context and situation;
- to enable rich communication media
- to provide awareness information.

The "smart awareness assistant" using information collated from sources such as on online calendar, your geographic location your expressed mood and any detected current activity could be used to encourage family members to share key pieces of information with each other and also to filter

out/prevent unwanted information exchange when the user is not available to communicate. This scheme could enhance togetherness by balancing awareness and privacy.

Figure 24. The lobby interface

5.3.3 Communication needs for the TA2 system

In order to answer the question which types of devices are best suited for meeting family members' communication needs, two focus groups of 6 and 5 participants of 2 hours and 4 individual interviews of about 45 minutes have been conducted (Demey, On usage and communication needs for the TA2 system, 2010). The interviews gave us more insight into the needs and wishes for sharing information with family or close friends depending on the state of the display (private, shared, semi-public), the device (TV, PC, mobile, picture frame), and the situation (group communication, synchronous versus asynchronous communication). The main findings were:

- The mobile phone is a controversial device. It is used for functional communication mainly, but also for social communication by younger users, for one-to-one or one-to-many communication.
- Long phone calls are popular and attractive. There is a need for enrichment with voice, video, pictures, etc.
- Voice-over-IP (VoIP) is attractive because of the low price and its video possibilities, but has lots of technological problems.
- Photo frame is recognized as a notification device. It is an always-on device, recommended for elderly in particular.
- Computers are complex and personal devices. Though they are powerful, they are not the best candidate for a family communication tool. They are excellent for organizational tasks (e.g. emails, blogs, organizing pictures).
- The television is an excellent device for group communication. It is suitable for video chat, gaming and picture sharing. It is seen as a device used for relaxing, so there is a need for easy interaction.
- There is a need for control, privacy and simplicity. Users are afraid of the "big brother" effect and want to keep in control of the system at all times.
- Social media are already quite developed. Including current services and enriching them is better than creating yet another application.

5.3.4 Scenarios for Enhancing Social Communication

Scenarios around sensors have been evaluated addressing:

- Functional communication, awareness & playing games
- Inter-Generational Communication
- Non-Intrusive Awareness
- Sharing Experiences – Pattern Recognition
- Social Awareness
- Visualizing Communication

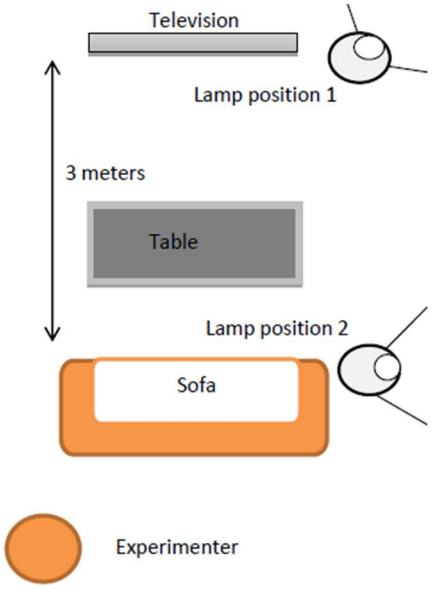
The consortium believed that the scenarios could be combined into an excellent TA2 application that combines smart toys (possibly as haptic devices), smart rooms (rooms containing objects enhanced with sensors), narrativity, real time communication and game-play. A clear target group would be that of grandparents remotely interacting with their grandchildren (with possible extensions).

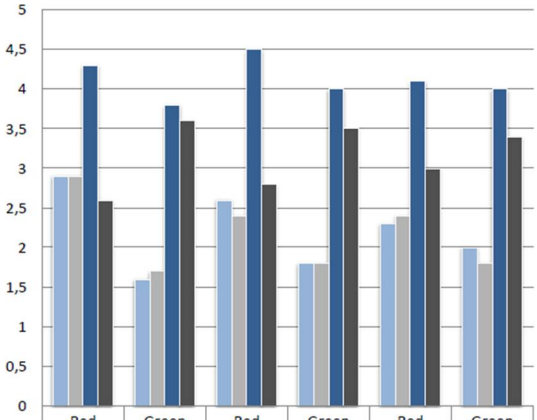
Grandparents love to keep an eye on children and tell them stories. Children love to hear their favourite stories over and over again. The TV is the main window of communication (audio-visual), but smart toys and smart objects act as natural input devices for interactive *Storytelling*. At the same time, smart objects and smart toys could become objects of attention, representing, for example, characters and places from the story, keeping thus the children more engaged into the story-world.

This led to the proposal of the *Storytelling*" demonstrator, described elsewhere in this document.

5.3.5 Enhancing communication using TV video chat and ambient lighting

In this study (Demey, Enhancing Family Communication Using Video Chat on TV and Ambient Lighting, 2011) we were interested in two main questions. First, we wanted to investigate the acceptance of television as a device for group-to-group video-communication in the home. This includes its influence on communication patterns within the family and between remotely located families, its effects on perceived social expectations or obligations, and whether it will raise any privacy issues. Secondly, since the television is not typically an always-on device, we were interested in using ambient lighting as a notification and awareness system.





	Red	Green	Red	Green	Red	Green
	Control condition		TV watching condition		Book reading condition	
I found the lighting annoying.	2,9	1,6	2,6	1,8	2,3	2
I found the lighting distracting.	2,9	1,7	2,4	1,8	2,4	1,8
I would like to be able to choose my own color.	4,3	3,8	4,5	4	4,1	4
I liked the color	2,6	3,6	2,8	3,5	3	3,4

Figure 25. Layout used for the lab test together with the results

To answer these questions, a lab study had been proposed for designing the lighting notification system. It served to better understand the possible usage of lighting as a notification system and how it is perceived by potential users. This lab experiment allowed us to calibrate the light notification for the field study (Figure 25).

After the lab test, a 4-week field study (Figure 26) with three households of a same family was run, where the television was enabled with a system allowing group-to-group video communication. The ambient lighting device provided information about the status of the system and also indicated when an incoming call was being received. This allowed us to test the validity of conclusion from the lab study in the field.

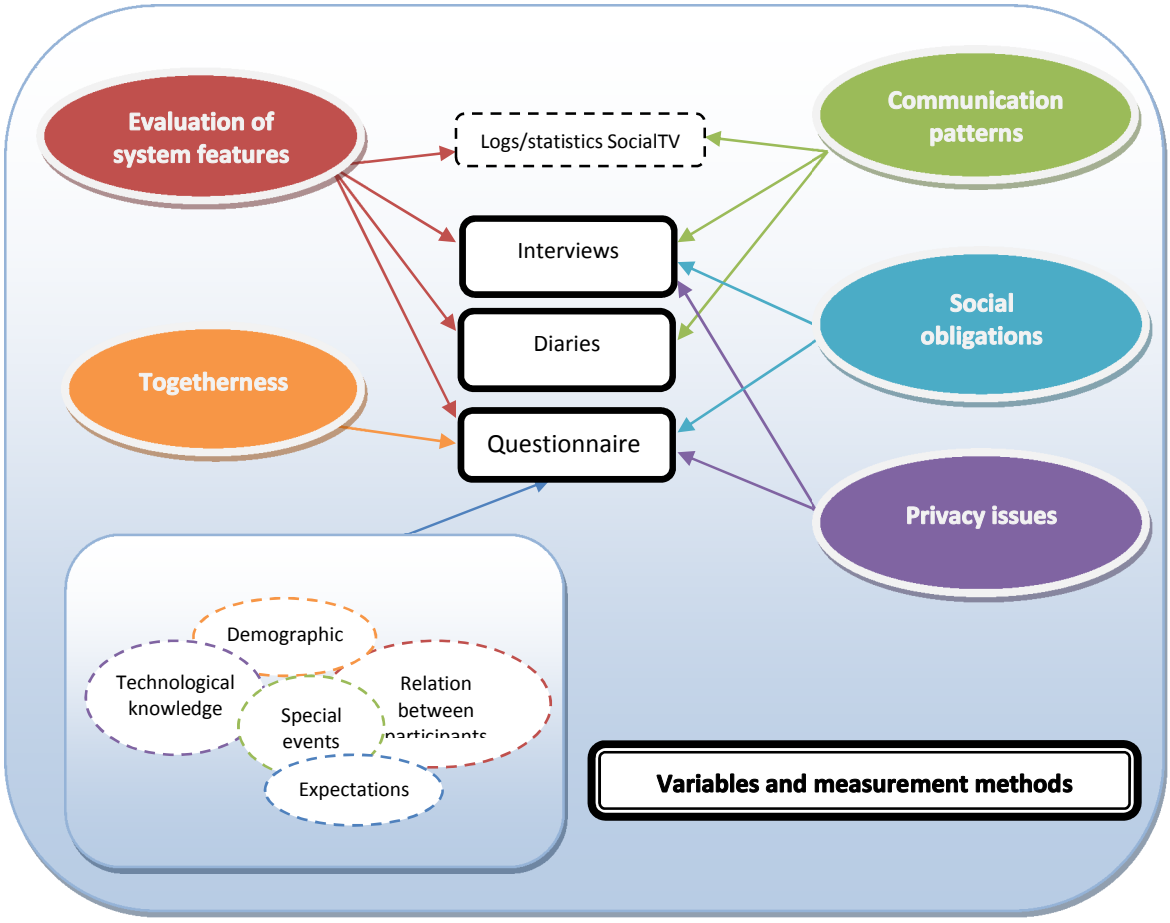


Figure 26. Variables (ellipses) and measurement tools (rounded boxes)

Some insights gained from the participants in the test:

- Video chat is not a replacement for the telephone call; it is another experience.
- Video chat requires 100% attention and presence (e.g. phone call allows for multitasking).
- Everything in the living room is heard and seen at the other end (visitors, children, clothes, furniture, etc.).
- Video chat is experienced as "going for a coffee", cozy and fun way of communication.
- Video chat is used to keep in touch, speak with children, show acquisitions.
- Video chat takes on average about 30 minutes.
- Ease of use: the simple UI in combination with air keyboard was valued as simple and easy to use.

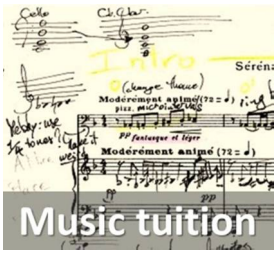
- Internet TV: Internet on TV is useful, faster than starting up the PC.
- Obligation: people where you normally visit unexpectedly are easy to contact, while for others you need to negotiate.
- Privacy: people behave and prepare like in a normal visit. Teenagers prefer to use their own PC for chat.
- Light notification is a very friendly way of passing information. It can be silently ignored, unlike a ringtone.

5.4 *The Connected Lobby* : What next?

The research around the *Connected Lobby* is feeding in to the thinking around Philips Net TV and in particular to the way the connected TV can be used as a vehicle for video chat. Future work will extend the ideas around the Smart Awareness System. We note that communication with families within the same time zone and with predictable schedules can be done relatively easily as leisure times are matched and an unplanned call, during what should be an available time, will most likely be received.

But if availability is scarce and if time zones reduce availability intersections further, actually having a video get-together may not be that easy; it may be necessary to negotiate an explicit time for synchronous communication. Picking a good time requires availability knowledge, but other's availability may not be predictable and availability information needs to be acquired to avoid guesswork. It is clear that when more than two families want to communicate, the complexity of the problem will increase. Understanding the effectiveness of a *Smart Awareness System* is a valuable next step and will be explored through further *in situ* studies.

6 *Music Tuition*



Description

A music lesson delivered over a video conferencing connection is an example of a shared activity (teaching music) in which the participants need excellent sound quality and the ability to see one another. With video the tutor can diagnose playing errors and difficulties based on what they can see. The tutor can also demonstrate correct technique and provide gesture-led encouragement and communication as the pupil plays.

Technology

The *Music Tuition* concept demonstrator adapts a commercially-available video conferencing system to work with some of the key capabilities developed within TA2. The additional functionality from the project includes the Video Router, which allows multiple cameras to be used and for seamless switching between them, and the Visual Composition Engine, which renders dynamic screen layouts defined in SMIL and allows control using a simple web interface.

Social Science

The relationship between a skilled musician and their tutor is critical. Trust and respect is essential and a social bond may be helpful. We believe that these can be supported and nurtured through an adapted video conferencing system. However, in this context success can be measured by a subjective proxy measure: did the participants find their lessons useful?

Evaluation

The *Music Tuition* demonstrator has been evaluated in a number of trials ranging from technical tests through to a series of lessons with six different orchestral instruments providing qualitative evaluations from a broad range of perspectives. Longitudinal tests for violin lessons between two sites in the UK will start in November 2011. Ethnographic observations of face-to-face lessons are being used to help identify ways in which the current implementation could be improved.

Learning

All participants see that video communication can provide a useful addition to face-to-face lessons when tutor and pupil cannot be together. Feedback highlights that:

- Sound quality dominates the perception of quality.
- Multiple camera views are perceived as useful but (so far) little value has been observed derived from the ability to dynamically switch from one view to another.

Ethnographic observations have highlighted that the musical score is a common point of focus (during face-to-face lessons) that should be represented in remote lessons.

TA2 experiences involve a shared activity and rich communication. Music tuition is a good example of a candidate TA2 experience as it combines a shared activity (the music lesson) with the need for excellent audio and the ability to see one another in order for the tutor to diagnose playing errors and difficulties (based on what the tutor can see) but also to allow the tutor to demonstrate correct technique and for the tutor to provide gesture-led encouragement and communication as the pupil plays.

The *Music Tuition* demonstrator offered the consortium an opportunity to co-develop with a renowned music school, Aldeburgh music, (see Figure 27) which had a need for a remote music tuition capability and who actively sought involvement with a technology project to fulfill this aim. The collaboration with Aldeburgh music thus brought to the project a live customer and the welcome opportunity to demonstrate how capability developed within the project could be integrated with commercial off-the-shelf video conferencing systems. BT is a significant reseller of video conferencing equipment in the UK and worldwide. This customer-led research into enhanced video conferencing design is being fed back into BT sales propositions teams. These teams will help to understand how insights gained can be used to either change the propositions we use with customers or to inspire feature changes with our suppliers or to direct subsequent work with system integrators with whom BT works in our video conferencing sales channels. Customer led research is powerful catalyst for exploitation as the work with the customer acts as a compelling case study for subsequent proposition development.

Figure 27 Interior view of the concert hall at Snape Maltings, the home of Aldeburgh Music

Aldeburgh Music (www.aldeburgh.co.uk) offers a year round programme of performing arts including the Aldeburgh festival, initiated by the composer Benjamin Britten (who was born near Aldeburgh) and his partner Peter Pears. Aldeburgh music is a place of energy and inspiration and provides opportunities for artists at all stages of their development to nurture and perfect their skills.

Aldeburgh is dedicated to developing artist talent. They were keen to understand how technology could help it do this whilst also controlling the need for travel with its associated cost and ecological impacts. Aldeburgh music sought solutions that would still be usable after the end of the project. This constraint brought with it the necessity to integrate some of the unique capability offered by TA2 technology with a fully supported video conferencing system; in our case a Polycom HDX6000 system. The project developed bespoke software and hardware solutions that brought additional functionality based on multi camera technology. The hybrid system, installed at Aldeburgh allows simple video conferencing connections to be made to any location with a compatible system; the more advanced capability is enabled by a simple web based user interface which can be activated from any remote site.

The desired uses of the setup are:

- *Music Tuition* one-to-one between student and tutor
- *Music Tuition* one-to-many between a tutor and a small ensemble such as a string quartet
- auditioning students who are geographically remote from Aldeburgh

Aldeburgh Music wanted a system that would enable them to better fulfill their purpose (auditions, master classes and artists development) even when students and artists are separated by great distances. We believe the following attributes could contribute to this goal:

- The ability to increase the frequency with which tuition can be offered.
- The ability to enhance the effectiveness of teaching by enabling lessons to be recorded and subsequently reviewed lessons in order to consolidate learning.
- Improved reach of Aldeburgh Music as a global brand in artist development achieved by the system reducing the impact of Aldeburgh's physical isolation.
- Improved artist development by providing student visitors access to great tutors even when those tutors were not able to be at Aldeburgh in person
- Improved artist development by offering vicarious learning by providing students with access to a library stored mater classes.
- The ability to discover, and to nurture, talent wherever it may be found in the world at lower cost, by facilitating auditions over the video conferencing system.

6.1 *Music Tuition*: the value of this demonstrator

Music Tuition enables the project to conduct longitudinal in-situ trials. It enables us to probe the value of aspects of the video chain, in particular the composition capabilities, and to understand how the system performs for slightly more formal relationships (that between a tutor and a pupil) rather than just for social interactions. This is highlighted in Figure 1.

Music Tuition is useful for evaluating the overarching research question asked by TA2 is: *How can technology help nurture social relationships between family and close friends?* Although it does not involve an interaction between family and friends (tutors and pupils have a slightly formal relationship though some level of trust and empathy is essential) the demonstrator is useful for the following reasons:

- **Design**: The involvement of Aldeburgh music in defining the use case and in quickly evaluating technology implementations has helped us to iterate our developments more purposefully to include the functionalities that are most likely to be useful.

- **Evaluation**: The installation of the system at the premises of Aldeburgh Music provides the project with the opportunity to observe and evaluate the system through in situ trials. Such trials provide the best evidence of whether the capabilities offered by the innovations we have developed are being used and (by inference) valued by the target users. They also provide a high level design goal: To provide Aldeburgh Music with a mechanism for effectively developing artist talent when tutor and pupil are not co-located.
- **Component re-use**: The demand that the components we develop integrate with an existing system helps our exploitation plan by giving the project evidence of how individual capabilities, outside the overall TA2 system can be integrated in order to bring additional functionality.
- **Exploitation and Dissemination**: The association with a world leading brand in artist development like Aldeburgh Music helps in dissemination and exploitation as Aldeburgh Music has stature within this field of endeavour and the way it chooses to reinvent its offer to artists will influence the thinking of other institutions.

The *Music Tuition* concept demonstrator involves performance tuition as a shared activity. Unlike the other demonstrators, this shared activity is not just about fun and communications but also about education of performance for professional level orchestral players. The *Music Tuition* concept demonstrator will augment existing music lessons not replace them, but will provide additional opportunities to receive tuition when separated from your tutor.

The objective is to leave a version of the *Music Tuition* demonstrator in use within Aldeburgh Music at the end of the TA2 project, this will provide a proof of concept for this application but also for certain components of the *Music Tuition* demonstrator that have been developed in the course of the project. Cooperating with a real-life customer from the beginning of the development has proven to be more than helpful in creating the application. Valuable input and feedbacks from the team at Aldeburgh Music helped in creating software that has been tailored to satisfy the specific needs of music teachers and pupils on an advanced and expert level who are spatially separated.

Developing the application in cooperation with Aldeburgh Music provides exploitation opportunities within Aldeburgh Music but also the prospect of developing for other music schools and institutions.

6.2 *Music Tuition*: concept development

Aldeburgh Music contacted BT in the summer of 2008 asking if there was a way in which a remote site like Snape (where Aldeburgh Music is based) could use video to improve their artist-development programme. This resulted in a piece of ad hoc consultancy (provide by BT outside the TA2 project) in which it became clear that Aldeburgh Music wanted to enable a distributed world orchestra. This ambition was found to be unfeasible given the annoying limitation provided by the speed of light; the transmission delays between Snape and potential site in Africa, America, China and India are (at 30-100ms) already sufficient to make playing traditional orchestral pieces together impossible.

Aldeburgh accepted the limitation imposed by physics but remained keen to understand the value that video conferencing could bring, even if ensemble playing was not possible. In the late summer of 2009 BT drafted a paper with Aldeburgh outlining some ideas on how video conferencing could be used. This included ideas being developed in TA2 such as multiple camera views, orchestration, and the ability to record and then interactively review the multiple views captured (much like the *MyVideos* demonstrator).

The possibility of involving Aldeburgh Music as a demonstrator within TA2 was discussed as early as June 2008 but was finally ratified in late October 2009. Following this discussion a number of use case documents were developed and rapidly iterated with Aldeburgh Music. Regular discussions on

the pros and cons of the different setups took place between representatives from the music school and the TA2 consortium. Eventually, the use case of connecting pupils at Aldeburgh with teachers in external locations emerged as the concept favored by Aldeburgh.

This was followed by technical design phase which sketched out how the capabilities could be integrated with established video conferencing equipment. The system architecture was discussed and developed through technical workshops. Subsequently, the components from the TA2-framework that were needed to modify and enhance the video conferencing system had to be identified. One central idea of the chosen concept is using several cameras in order to offer the tutors different views on the pupils. The TA2-framework supports multiple cameras in one location, and a video composition engine (VCE) has been developed to handle multiple video streams. The idea was to let the main camera be controlled by the video conferencing system and the side cameras by the VCE. The main challenge consisted of merging, arranging and controlling these video streams on one screen.

The system architecture agreed to deliver this capability is shown in Figure 28

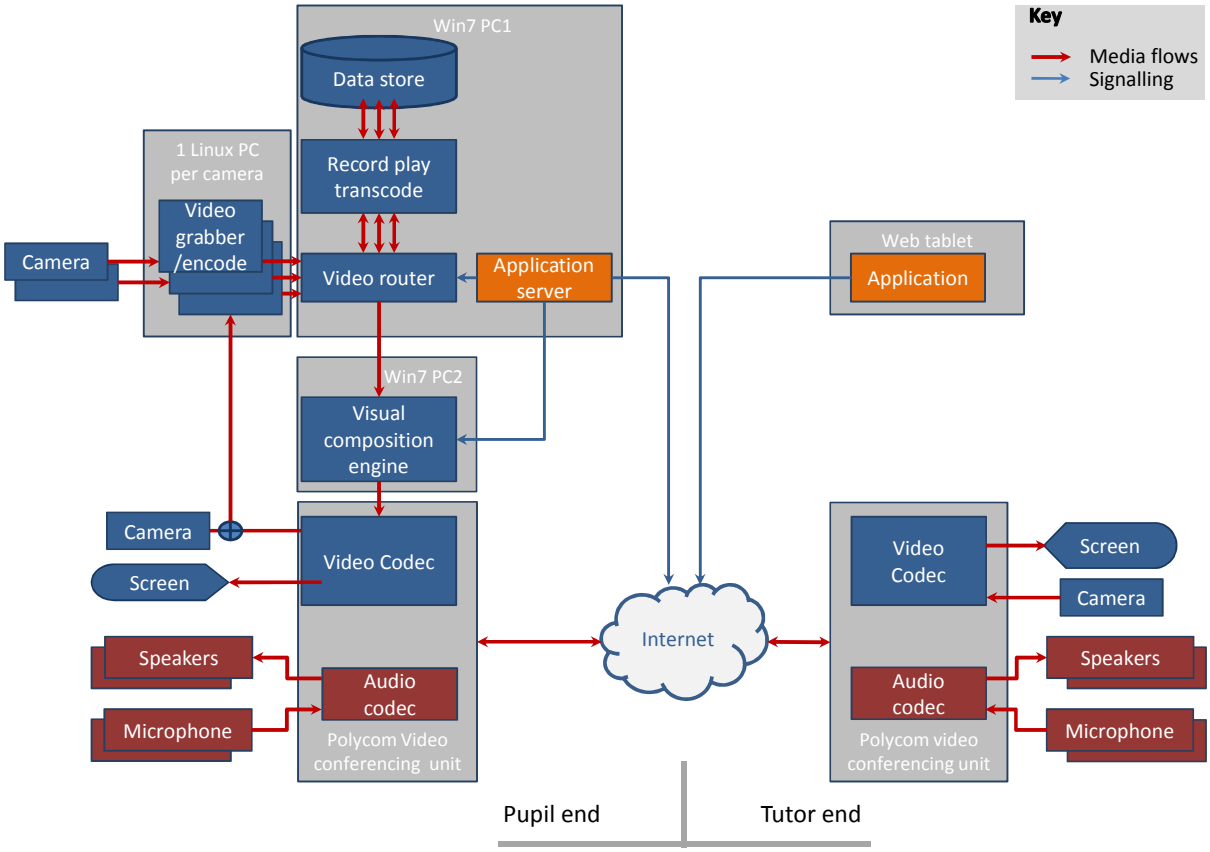


Figure 28 System architecture agreed to deliver the *Music Tuition* demonstrator

Choosing the video conferencing system was a process which involved many parameters and took place in summer and autumn of 2010. There were technical, qualitative and financial constraints, so the process of defining the manufacturer and the specific system involved many tests and considerations. The Polycom HDX6000 emerged as the system of choice. One of the main reasons for this was that Polycom invoke a particular mode for audio transmission that is well suited to music (Music Mode). This involves a high sample- rate codec with minimal audio processing. Music mode is a system profile that has been developed in conjunction with Manhattan School of Music specifically for remote tuition and performance.

A further challenge was providing sufficient network bandwidth for the system to run well. (This is an excellent example of a challenge that is only recognized when working with a real customers). The network solution provide by a company called *Sharedband* bonds four ADSL broadband lines so they appear and behave as a single logical connection with (in principle) four times the overall bandwidth of any one connection.

In spring and early summer of 2011, system integration began. The first system prototypes were produced and internally tested. After the basic functions were available, a series of technical pre-tests were held from March to June 2011.Testing was successful and identified weaknesses in the software and hardware setup as well as in the usability of the interface. Most of the technical issues were addressed and some of the user input was used to improve the system before the evaluations proceeded.

A first set of 1:1 end user evaluations took place in July 2011 at Aldeburgh. The evaluation results are presented in the next chapter. An outlook on further plans and upcoming developments of the *Music Tuition* concept demonstrator are described in the last chapter.

6.3 *Music Tuition*: Evaluation results

Several evaluations have taken place and more are planned - the whole process is iterative. The initial tests were largely technical. These were followed by a series of tests with musicians and will lead to longitudinal tests between tutor pupil pairings over several months.

6.3.1 Connection test: Test of the connection between the Manhattan School of Music (MSM) in New York and Aldeburgh Music

The setup included two Polycom systems, (an HDX4000 at the Aldeburgh end); both augmented with external speakers. The session was held by a Cello lecturer who was visiting Aldeburgh, and who works with the MSM distance learning program on a regular basis. The tutor was giving a lesson to one of his regular pupils located at the MSM.

Figure 29 Photographs of the Connection Test in which the network performance of the Aldeburgh Music set up was tested

The connection between the two Music schools was successfully established. All the components worked, the received pictures were HD quality and the audio was clear. Some instability (frames being dropped and pictures freezing momentarily) was corrected by changing router settings in order to reduce jitter. The general impression of the participants was positive. After the test the proposed extensions to the system (multiple cameras/views and a separate hand held input device and a better sound setup) were discussed. Perhaps most critically the student really valued the lesson as they were

preparing for an imminent recital; they were convinced that the input from the tutor had helped improve their performance.

6.3.2 Software test: This test was a technical test to verify that the hard- and software setup was rugged enough for the first user tests

The purpose of this test was to build confidence in the multiple camera system and the overall technical setup and to introduce the system setup to the technical staff at Aldeburgh Music.

The tests were conducted with two cellists, one located at Adastral Park, the premises of BT, and the other at Snape, the home of Aldeburgh music. The tests revealed some issues with the hardware setup, as well as with the synchronicity and usability of the software components. The main issues were:

- The lighting conditions in the room were problematic for the installed cameras
- The image fidelity from all but the central camera was not yet good enough
- The audio quality was not good enough to diagnose and appreciate subtle features of play as required by players of this calibre.
- There were synchronisation issues (a delay between audio and video that built up over time)
- The views presented on screen were sometimes disconcerting (the participants didn't want to see themselves whilst playing) and viewed the layouts (with both ends having the same set of views all the time) as being unnatural.

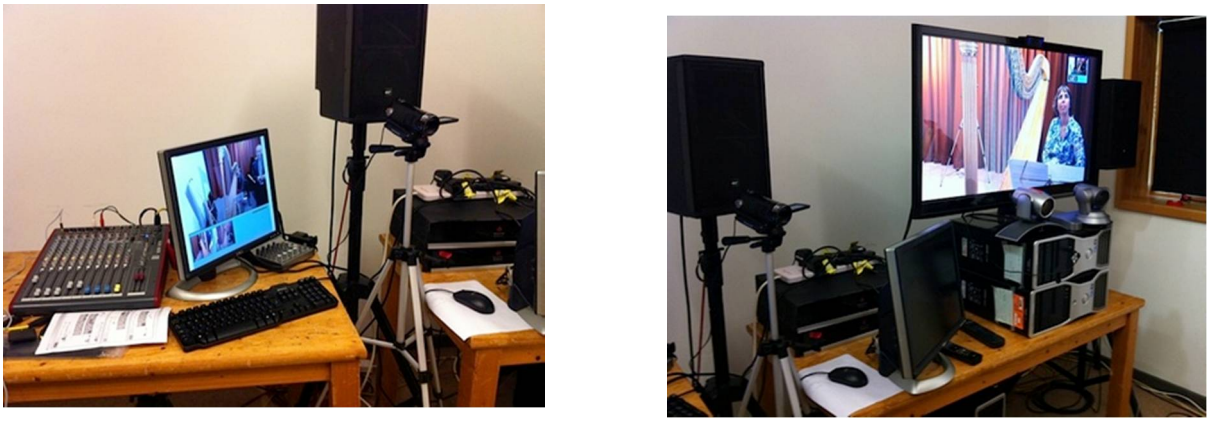


Figure 30 Photographs of the system set ups used in the tests with musicians at Aldeburgh Music

The overall impression of all people involved was very positive. The tablet screen switching worked well with the system switching gracefully between different camera views. The test enabled us to define camera requirements for the multiple camera system and the technical components needed to run the system at Aldeburgh independently were identified. Criticisms of the audio and video quality were substantially addressed through deploying different hardware (Cameras and loudspeakers) and none of the other issues were so problematic that we could not solve or mend them before the first user test in July.

6.3.3 First user tests: Observations of a range of music lessons using different orchestral instruments

The objectives of this evaluation were:

- Observation of at least three lessons
- Pupil interviews
- Tutor interviews
- Experience of setting up cameras for three different musical instruments

In addition to new speakers and cameras (both of much higher quality than those used in the technical test) an additional screen was added in the VC-room, so two camera views could be presented full screen and in high resolution (one showing the tutor and the other showing the pupils as they could be seen by the tutor). In order to address the synchronization issues in the previous test and to keep audio and video in sync, the audio delay could be manually adjusted to match the much higher video delay. Some of the software components were also updated to provide more stability and a better view-switching experience.

The test took place in two rooms at Aldeburgh. The system setup in the pupil room included the modified Polycom video conferencing system, one main and two side cameras, two HD-screens, a set of external speakers and the standard microphone built into the Polycom system.

Violin	**Harp**
• (Tutor) had participated in our tests before and stated that the sound quality had improved and the music sounded better than last time • There were concerns about the dynamic range, but since tutor and pupil did not know each other before, it remained unclear whether this was due to the system or the way the pupil interpreted the piece • The tutor used the tablet to select different camera views only once or twice • (Tutor) was looking at the three views but didn't feel the need to bring one of them into focus • (Tutor) might have found the ability to zoom useful	• Issues with lighting and reflections made it difficult to see the strings • Being able to see the pupil's feet and hands is essential for teaching this instrument • The ability to see both hands/both sides of the instrument is helpful • Used tablet to change views • Lesson was an interactive experience
French horn	**Piano**
• The room was criticised for not offering a good sound stage for this instrument • Pupil and tutor experienced acoustic issues - presumably with the echo cancellation • Horns are right/left-handed, so orientation was an issue. • Participants need to look into the bell of the horn in order to see stopping actions • General comments were positive and the tuition was considered viable	• The camera view showing hands and feet was considered useful • The audio delay was not right when the lesson started, but this could be adjusted • Loud passages caused massive sound distortions – this problem could not get fixed during the session, neither by repositioning the microphone nor by modifying the input level settings
Cello	**Oboe**
• Had some issues with light reflexions on the instrument • Experienced some distortions at certain sound ranges • Noticed the visual delay • Did not use the tablet at all • General impression was that this was a very effective lesson	• Experienced some sound distortions • Tried to play together/synchronously - this was not possible due to the delay of the system • Was not comfortable with the tablet and did not change views

Table 1 Summary feedback from the six lessons.

The tutor's room was equipped with an unmodified Polycom video conferencing system. Both sound and video were transmitted by the Polycom system. A touch screen control panel was provided which allowed the tutors to control the video input in the pupil's room thus allowing the tutors to choose which of the views of the pupil's room studio was to be shown in the main view of their screen, see Figure 31.

The participants for this evaluation were recruited from the Aldeburgh Young Musicians programme. Six lessons were organised, each featuring a different instrument. Major comments and observations from the six sessions are summarized in Table 1.

Overall, the evaluation has been very successful and all the initial objectives have been met. Six lessons were held, observed, and later discussed with the participants (instead of the originally planned three lessons). The team working on the application and observing the evaluations gained valuable insights on the dynamics of music lessons and on the specific demands concerning the camera setups for different types of instruments. Most of the feedback given in the interviews was very supportive of the concept and contained many constructive details and suggestions. Especially pleasant/encouraging was the feedback of the principal of Aldeburgh Music, Jonathan Reekie. He stated that he definitely sees potential for remote learning based on the systems developed so far.

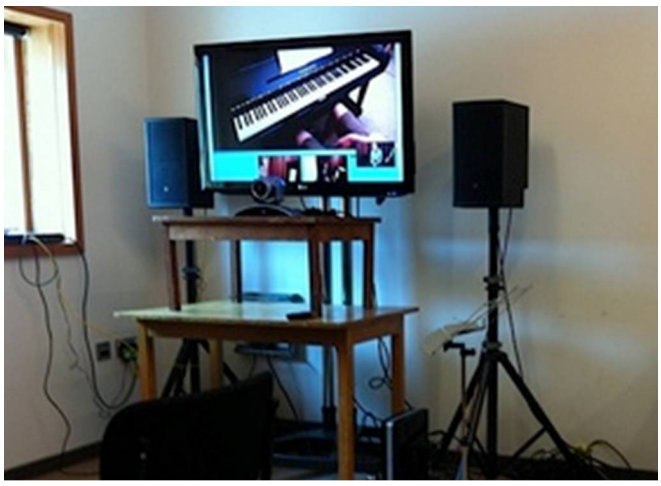


Figure 31 The screen view available to the tutor. The screen layout shows a large picture in view and three other options through which they can toggle using a touch screen device (not shown)

6.4 *Music Tuition*: What next?

The evaluations provide much support for the assertion that the augmented video conferencing system will help Aldeburgh Music to better fulfill their purpose (auditions, master classes and artists development) even when students and artists are separated by great distances. But, they have also identified many parts that could be improved, extended or added.

Ease of use: An admin interface is going to be created in order to make the system robust and operational. In case of a critical error, a system shut-down or a loss if internet connection, this additional interface should enable users to reboot the system and restart the application without the help of an admin or other technical staff. Providing an admin interface is highly prioritized because of the importance to grant Aldeburgh the ability to use the system according to their plans and without further assistance from the TA2-team.

Sound quality: The evaluations in July showed that under certain circumstances the audio was distorted clipped or poorly encoded. We expect to address this by improving the analogue capture of the audio and bv introducing limited sound processing at this input stage. This sounds simple but may highlight a limitation of the Video Conferencing system we have chosen. The HDX60000 is only able to accept digital audio input. We are working with Polycom to see if there are any workarounds possible.

Use case scope: Aldeburgh Music would like to know whether the system can be extended to also support auditions. This would require a record and review functionality. As a first and most important step, the recording of whole sessions is requested. Pupils will then be able to audition in designated places which offer access to a video conferencing system, so it will not be necessary for them to travel to Aldeburgh.

Sharing the score: Ethnographic observations of music lessons when tutor and pupil are in the same room has highlighted that the score is a critical artifact through which the tutor and pupil communicate. Presenting and sharing the score of the currently played music is another feature that has been requested by Aldeburgh Music. This request is fairly new and had not been considered in the design of the previous use cases. Suggestions will be developed in cooperation with Aldeburgh in order to get a better understanding of the purpose and aims of this feature. Potential solutions will then be presented to the technical team and it will be determined whether the expected results will be worth the effort of implementing this element.

Network performance: Aldeburgh music would like to see the Shared Band solution they are employing achieve its potential throughput. At the moment, for system design reasons within BT's network the throughput is limited. There are potential ways around this but we also look forward to the fibre enabling of the local exchange which should provide higher bandwidths through the existing lines.

User interface enhancements: The interface on the control device is a work in progress and several minor display issues are being looked into. One result of the evaluations showed that the interaction with the touch screen control device conflicts with the ability to use an instrument. In order to enable tutors to demonstrate playing an instrument and to simultaneously interact with the system (especially switching the video views), a foot pedal input was suggested. Adding an input device based on a foot pedal could offer a hands free way for tutors to toggle the view. The hardware is available and cheap and the costs to integrate a foot pedal as an alternative input device seems well worth the effort.

During the Autumn of 2012 the system further tests will be carried out between a pupil and student pairings between London and Aldeburgh. This will provide inputs on whether efforts to address some of the hygiene issues (interfaces, usability and sound quality) have been effective. Ethnographic observations of these sessions are planned and should help us to understand the limitations of lessons held remotely and hopefully provide the starting point for further developments.

By the spring of 2012, the system should be running as a stable system requiring little input from the TA2 team.

7 Summary and Conclusions

Concept demonstrators have been used throughout the course of this project to assess how technology affects the way the relationships between close family and friends can be nurtured. An iterative approach has been used to select and develop the demonstrators and this has resulted in the nature of the demonstrators changing during the course of the project.

The concept demonstrators have been selected to allow variety within the following parameters

- The technology used within each demonstrator.
- The nature of the demonstrators in terms of the number of endpoints whether the interactions are synchronous etc.
- The age of, and the nature of the relationships between, the participants.

The demonstrators active during the final eighteen months of the project include:

- *Family Game* – which brings the experience of a family board game to people in separated households.
- *Music Tuition* – a system testing the value of multiple camera set ups in remote music tuition.
- *Storytelling* – a means for allowing a bedtime story to be shared between different households
- *MyVideos* – a system to allow the development of personalised videos collated from content captured at a school concert by many of the audience.
- The *Connected Lobby* - a means by which households engage in any of the above activities based on presence technologies displayed through a TV.

Evaluation of these demonstrators is still going on but it should be anticipated that the evaluations will provide qualitative evidence that supports the notion that technology capabilities implemented within shared applications will positively affect the way relationships are nurtured.

The following are interim selected conclusions from the work carried out to date.

Family Game

- Whilst Game playing is an important activity for groups and does help to build relationships, gamers report that tactile elements of the game (the board, the pieces etc.) are important aspects that should, if possible, be retained.
- Game design should focus on game play in which communication is required; cooperative games are likely to be a good choice in this regard.
- In order to enhance the value derived from the communication, the game and video elements should be composited on the same screen to encourage eye contact. This requirement will emphasise capabilities of the visual composition engine component developed in TA2 and is being tested in our trials scheduled for Nov 2011.

Music Tuition

- The system developed by TA2 enabled music lessons that according to the pupils and tutors are effective and according to some comments “just like a normal lesson”
- The use of multiple camera views appeared important for some instrument types but not all.
- Ethnographic observations of real face to face lessons highlight the importance of the physical artefact the score during lessons. We anticipate the systems would be more useful if an

effective way of sharing the paper based annotated score, as used by the pupil, could be devised.

- The standard audio capture provided by the Polycom system used in the experiments lacked the dynamic range to effectively encode loud piano and horn pieces; improved audio capture based on analogue and mixing desk like capabilities will be an essential component of improved music tuition systems.

Storytelling

- The ability to perform evaluations within people's home provides insights into the domestication of technology that cannot be achieved in the laboratory and that cannot be easily anticipated. Such experiments, whilst a compromise in terms of assessing 'bleeding edge' technology, remain extremely valuable.
- Quantitative data on the number of times books are read, the time spent on each page whether books are re read and whether participants utilise the interactive elements is being collected and will be available in late 2011/early 2012
- Users have invented their own use case for shared applications which require flexibility in the placement of microphones and cameras in order to facilitate a particular kind of sharing (a view on a crossword puzzle) or to improve the fidelity of the interaction (by moving the microphone closer to the speaker). Future designs of such system should explore the use of mobile video and capture devices and to consider whether current models for spatialising audio and orchestrating the video are robust enough to accommodate capture devices that can move.
- All users without exception have asked whether they can keep the TA2 setup. This is an encouraging qualitative indicator of the value users perceive in the technology.

MyVideos

- Users believe that systems like *MyVideos* would encourage them to capture more video material and to share more video. Assuming that the sharing of stories is an important part of building togetherness this would suggest that the ability to generate personalised stories from shared content will help nurture relationships between people that know each other well.
- Users of the early *MyVideos* system were (generally) pleased with the personalised presentations that they produced with the system.

Connected Lobby

- The television is seen as an excellent device for group communication with associations to social and relaxation oriented behaviour.
- Video communications through the TV is seen as an addition to, and not a replacement for, telephone calls. One user referred to it as being more of an event – "like going for a cup of coffee". This perception is useful in understanding how users will incorporate Video Communications into their lives and therefore an important insight to be used in presenting, through marketing message for instance, how such products could be used.
- User have varying perceptions of the privacy threat posed by microphones and cameras within the home. An effective way of dealing with this relative uncertainty is to provide users with clear control over these capture devices.

- Initial tests on the use of ambient lights that turn on to indicate presence or flash to indicate an incoming request for a video call suggest that green lights are less distracting /annoying the red lights; however responses were even more positive to the notion that users could choose the colour of the light notification.
- Generally, lights are considered a friendly way of representing such information. Users commented that the flashing light could be silently ignored unlike a ringtone and viewed this as a positive feature.

Using concept demonstrators as a vehicle for evaluating the usefulness / value of technology is powerful, but if the project is approached with a rigid mindset and with a focus on a single technology development rather than a high order goal it can be a little frustrating. Many of the valuable insights from the user interactions have served to identify important considerations that were not necessarily primarily dependent on the technology innovations upon which the technology teams expected to work:

- *Storytelling* has highlighted issues about form factors for microphones and camera housings and has no yielded significant deep insight into the value of spatialising audio. Tantalizingly, every triallist wanted to keep the system; is this because they really value the connection it provides to a close friend – or do they just like the idea of keeping the iPad? These are questions that we will seek to answer in the next few months.
- *Music Tuition* has highlighted the need to share the score in order to make music lessons more natural and the need for better audio capture in order to manage the large dynamic range of instruments like the piano forte, but insights on the use of multiple cameras (as we imagined it) are, as yet, mixed. However the observed requirement re the score suggests that multiple active cameras and (possibly) intelligent composition of these views will be important.
- *Family Game* has revealed something of the importance of the way video and media elements are composited but has also helped us understand the importance of game choice and of the human value of physical artifacts in game play.
- The *Connected Lobby* and early trials with TV based video conferencing suggest that a video chat on the TV is like 'going for a cup of coffee' and not a replacement for a telephone call. This hints at an intimate relaxed interaction that is really valued – but will users invest the time in such interactions when the novelty of having the capability has worm off? At the same time practical implementation experiments have highlighted the relatively paucity of support for real time communications within the current breed of Connected TVs.
- *MyVideos* showed that people enjoyed the ability to create personalized stories and that they thought the system created pleasing and appropriate stories – but would they then be more likely to share them and would this sharing effectively nurture the relationships.

The demonstrators used within TA2 have generated a lot of evidence suggesting, by inference, that technology can help nurture social relationships between groups of people that know each other well by enabling the groups to enjoy a shared activity.

Some of the conclusions can be related to the technology capabilities developed within the project. *Music Tuition* and *Storytelling* both revealed users demands for roving cameras that could capture additional shared content (A musical score and a crossword puzzle were the particular examples). We infer from this that the ability to, intelligently, compose on a shared screen images involving both the view of the remote person and the additional camera are valued. The ability to capture transmit and intelligently compose multiple views is one of the key developments within TA2.

In use, the technology is often not the headline factor that affects the value of such interactions. It is absolutely necessary, a hygiene factor, but value, from a users' perspective is often driven by the things they can affect – the physical nature of the set up or the nature of the interaction, not by the complex but invisible technology that supports it all. But then again few of us acknowledge on a daily basis the value that oxygen brings to our lives. It doesn't mean it isn't important.

8 Works Cited

Ames, M. G. (2010). Making Love in the Network Closet: The Benefits and Work of Family Videochat. . Savannah Georgia: ACM Press.

Chen, X. (2009). *[4] Chen, X. (2009). Strategic Design for Connected Lobby to Enhance Family Togetherness by Balancing Awareness and Privacy.* Delft: Delft University of Technology, Delft, the Netherlands.

Collins, R. (2005). *Interaction Ritual Chains.* Princeton University Press.

Demey, C. (2010). *On usage and communication needs for the TA2 system.* Eindhoven, the Netherlands.: Eindhoven University of Technology.

Demey, C. (2011). *Enhancing Family Communication Using Video Chat on TV and Ambient Lighting.* Eindhoven, the Netherlands.: Eindhoven University of Technology.

Dourish, P. (2001). *Where the action is: The Foundatiosn of Embodied interaction.* Cambridge Massachusettes: MIT Press.

Ducheneaut, N. M. (2008.). Social TV: Designing for distributed, sociable television viewing. *International Journal of Human-Computer Interaction*, 24 (2), 136-154.

Durkheim, E. (1971). *The elementary forms of the religious life.* Allen and Unwin.

Eichhorn, E. (2008). *Thoughts on Social Communication.* Eindhoven, the Netherlands.: Philips Consumer Lifestyle.

Geerts, D. (2006). Comparing voice chat and text chat in a communication tool for interactive television. *Proceedings of the 4th Nordic Conference on Human-Computer Interaction 2006: Changing Roles. NordiCHI '06, vol. 189.* New York: ACM Press.

Goffman, E. (1967). *Interaction Ritual: Essays on Face-to-Face Behavior.* Anchor Books.

Harboe, G. M. (2008). The Uses of Social Television. *Computers in Entertainment, 6, 1.*

Huang, E. H. (2009). Social Television Comes Home: A Field Study of Communication Choices and Practices in TV-Based Text and Voice Chat. *Proceedings of CHI 2009.* ACM Press.

Kock, N. (2004). The Psychobiological Model: Towards a New Theory of Computer-Mediated Communication. *ORGANIZATION SCIENCE*, 327-348.

Ling, R. (2008). *New Tech, New Ties: How Mobile Communication Is Reshaping Social Cohesion.* MIT Press.

Ljungstrand, P. a. (2008). Supporting group relationships in mediated domestic environments. *Proceedings of Mindtrek 2008* (pp. 59-63). Tampere, Finalnd: ACM Press.

Williams, D. U. (2011). Video mediated social interaction between groups: design guidelines and technology challenges. *Telematics and Informatics Vol 28, No. 4.*, 251-270.