“Powerpoint to the People”: Suiting the Word to the Audience

René Hexel **Chris Johnson** **Bob Kummerfeld** **Aaron Quigley**

School of Computing and Information Technology
Griffith University
Nathan Campus, Australia
r.hexel@griffith.edu.au

Department of Computer Science
Australian National University,
Canberra, Australia
Chris.Johnson@cs.anu.edu.au

Smart Internet Technology Research Group
School of Information Technologies,
University of Sydney, Australia
{bob,aquigley}@it.usyd.edu.au

Abstract

A computerised system supporting public presentations that are “personalised” at two levels is now possible. Firstly, the system exploits context information to adapt the large-screen projected presentation on the basis of who is in the audience. Secondly, the system makes use of the display devices of individuals in the audience: these provide an additional and complementary display with both content and presentation adapted to the individual. We describe the architecture and issues in supporting this level of personalisation.

Keywords: Intelligent Context-Aware Environments, Adaptive User Interfaces.

1 Introduction

Presenting a talk to a small audience is a familiar scene of everyday computer technology. The speaker has an audience, a personal laptop computer containing a presentation, and a data projector. The speaker’s talk is accompanied by a series of slides containing words and graphics, video and audio clips, projected onto a screen visible to all. Usually the speaker has a separate monitor that mimics what is being projected on the main screen, or she has her laptop screen as a monitor. Newer laptops and presentation programs allow an alternative view of her presentation to be displayed on the monitor, perhaps containing the overview of the talk or her notes associated with each slide.

The audience sees the same slide show, on the common screen, even though many of them have their own PDA or laptop that is capable of displaying a view of the slides. The members of the audience may speak many languages. They may have many affiliations: different departments, allied enterprises, other companies. The slides for this particular presentation are chosen by the presenter from what she knows or guesses about the audiences’ interests and affiliations: technical, managerial or financial aspects of the topic, confidential or public versions. The language of the words on the slides is chosen by the presenter.

The speaker’s laptop controls the presentation: by automatic timer, or her use of keyboard, mouse, or remote, she attempts to ensure that the projector’s progress through the slides matches her progress through the talk, and all members of the audience see and hear the two in synchrony under her control. This synchronisation is a desirable feature of this kind of presentation, the shared simultaneous experience allowing the people in the audience to associate the words and graphics on the slides with the speaker’s verbal explanations, reinforcing their understanding and memory. Questions and discussion from people in the audience during the presentation are also synchronised with the presentation, and can further contribute to the shared experience. But there is only a single, common style and rendering of the presented slides, necessarily set to the assumed lowest common level of the audience and the presentation technology. What is missing is any personalisation of the shared presentation: adapting the presented materials in some way for the individuals and their own presentation devices, and enabling the content of the common presentation to be more easily tailored to the roles, affiliations and interests of the audience.

In the next section, we outline how the quality of a presentation could be significantly enhanced by personalisation. We describe a presentation scenario and the technology for context awareness and the current challenges for such a scenario. In Section 4 we propose an architecture for personalised presentations and show how this architecture can be used to achieve the goals outlined in this proposal. In Section 5 we discuss the infrastructure requirements and implications of our architecture and present our conclusions.

2 A Personalised Approach

With a personalised presentation, people with less ability to understand the language of the presentation could be assisted with translations of the text on the slides as supplements. At present, a thoughtful speaker can provide these as separate printed notes or Internet-accessible versions. But notes must be prepared and brought to the talk and it would be hard to anticipate the variety of preferences that might be needed for any particular presentation, even though versions of the talk might exist in many languages.

The video clips embedded in the presentation may exist with a variety of soundtracks and subtitles in various languages (or be translated to have them), but for individuals in a general audience, these require some auxiliary, private display: the video content of the clip would be obscured if all subtitles were projected simultaneously for the needs of all potential audiences.

Similarly, people with varying degrees of visual or auditory acuity may benefit from larger fonts, colour-modified graphics, or frequency-selective boosted sound, but these require selective adaptation of the content to the individual display and sound systems.

An improved human interface for this situation is desirable. The elements of the required technology are at hand:

- widely deployed local wireless network cards in laptop and handheld computers, and widely accepted protocols for short-term, short-range connection to local networked services
  WiFi and DHCP, Bluetooth

- standardised or open representations of multimedia resources and presentations comprising multiple alternative versions (OpenOffice.org XML, MPEG-21, at varying degrees of complexity and standardisation)

- methods for selecting and adapting from such content (such as MPEG-21 Digital Item Adaptation or alternatively, XML representation, XSLT transformations defined in XSL, applied by filters deployed in versions of OpenOffice.org)

- standardised or open description of the preferences of "users" (our audience members)
  MPEG-21, UAPROF

- standardised description of capabilities of actual and virtual graphics display devices, and common occurrence of a few categories of displays, in similar windowed environments
  CC/PP, MPEG-21, UAPROF

- widely deployed tools for projecting dynamically changing bitmap graphics onto windowed displays
  VNC

- higher level distributed services
  such as the *Merino* architecture for distributed context services, which provides the higher layer of services and organisation to support the design of the personalised presentation (Kummerfeld, Quigley, Johnson & Hexel 2003)

The challenge we face is to engineer a combination of these technologies that adapts the computer interface to the human preferences and is easily deployed, publicly accessible, sufficiently secure and trustworthy. Its goal is to provide the presenter and the presenter's controlling application program with knowledge of the audience members' preferences, and provide audience members with individually tailored slides and content synchronised to the common slide show. We consider the issue of personalising the slide show to the audience members' preferences of language, video and audio, and to the presenter's choice of particular parts of the content, dynamically based on the audience members affiliations. Our Merino architecture for distributed context services provides the higher layer of services and organisation to support the design of the personalised presentation (Kummerfeld et al. 2003).

The Technology: Context-Awareness

This scenario can be seen as an instance of context-aware computing, making use of two of the four generally accepted aspects of context: the *where* and *who* (the others being the *when* and *what*, following Dey and Abowd (1999)). The context awareness is made possible by identifying the members of the audience sufficiently in order to determine their personal preferences, and describing the capabilities of their personal computing devices sufficiently in order to determine their display properties. This requires some identification of who is actually a member of the audience. This is not as simple as it sounds, if location is given in terms of geometric coordinates: unless the venue has a location proximity detector which has approximately the same bounds as the room in which the presentation is taking place, simple knowledge of the location coordinates of a number of people does not tell us which ones are audience members. Since we wish to intrude as little as possible, and use minimum additional instrumentation of people, we will use the personal computing devices themselves as the primary means of identification of the people that carry them. We may ask them to register their interest in the seminar by active registration with the room server, but we acknowledge that we run the risk of a lurker in the next room viewing our presentation without themself being visible.

Here is a more detailed description of the scenario. The venue is equipped with a computer that drives the local data projector, is connected to the Internet with a reasonably high bandwidth connection, and also has access to the local wireless LAN. An application running on this computer is the Presentation Context Manager (see Figure 1).

Tim Mansfield, a research scientist, has been invited to present at the University of Sydney an overview of his research and industrial effort in the "Information Ecology" project. Weeks before the talk, Sam the secretary enters a diary note for the presentation by Tim Mansfield on 12 February 2003 in seminar room G92, Madsen building. The diary note contains a pointer to Tim's home page.

Five minutes before the talk, the Presentation Context Manager at the venue turns on the projector and delivers a locally generated welcoming slide (derived from the diary entry, which is sufficient specification of the activity taking place in this room). As each member of the audience enters the room and registers their laptop or PDA with the Presentation Context Manager, their identities are determined from their wireless card's MAC addresses. The Presentation Context Manager retrieves their personal details (or at least a publicly visible persona containing sufficient details for this activity) from a remote service (Kay, Kummerfeld & Lauder 2002, Kay, Kummerfeld & Lauder 2003).

As Tim enters the room, the welcome slide is already displayed. His own Presentation Controller application on his PDA has already been activated due to the entry in the diary. The Presentation Context Manager has registered the PDA's presence on the wireless LAN, and the wireless card's MAC address identifies it as belonging to Tim Mansfield; the Manager knows him to be the presenter. The Presentation Context Manager provides the Presentation Controller with what it knows about the identities of the audience.

Tim will use this information to decide what slides might be excluded from the talk, according to his perception of the audience members affiliations and interests. His Presentation Controller asks the Presentation Context Manager to retrieve the contents of the presentation from his home server, or from its own disk. The Presentation Controller displays the presentation as a series of thumbnails and allows Tim to delete them selectively for this occasion, while he is still waiting for the last few audience members to arrive.

Two of the audience members have expressed a preference to use German subtitles for this presentation, so Tim is informed of this fact. He does not need to do anything further about it, because he has prepared the content of the presentation with alternative text in several languages.

Tim starts the presentation with his first slide. The first slide appears on the main display screen with a greeting naming the identifiable individuals in the seminar room. Versions of each slide are selected or created by the Presentation Manager, rendered at the server, and different versions are sent to the data projector for the main screen and others, suitably transformed for the graphic capabilities of their devices, for those audience members who are subscribing through their own laptops or PDAs. The al-

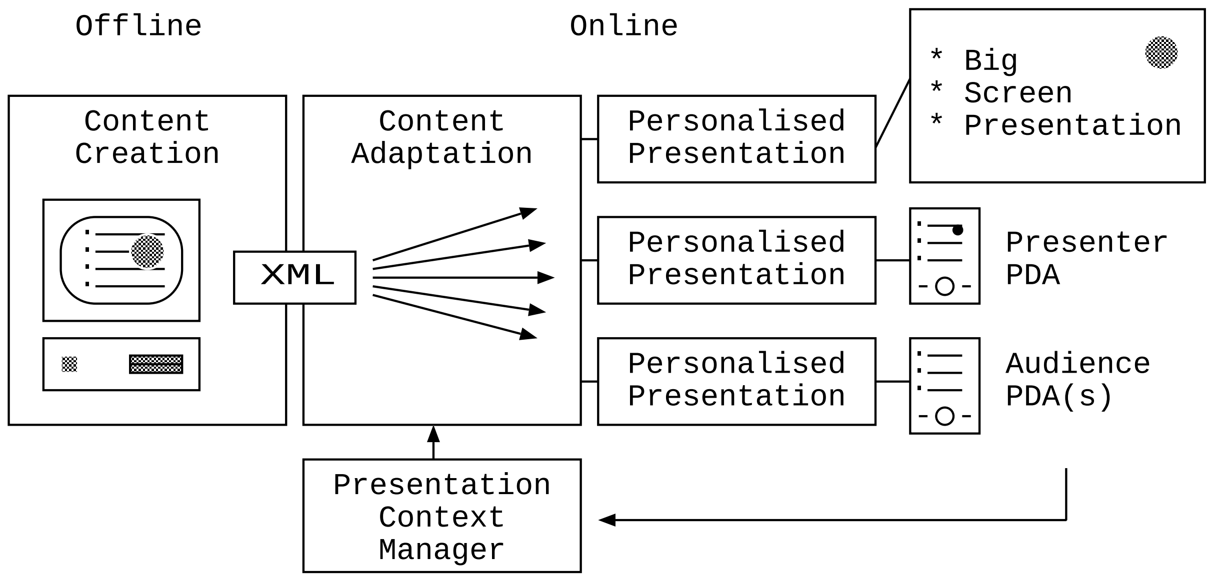

Figure 1: System workflow

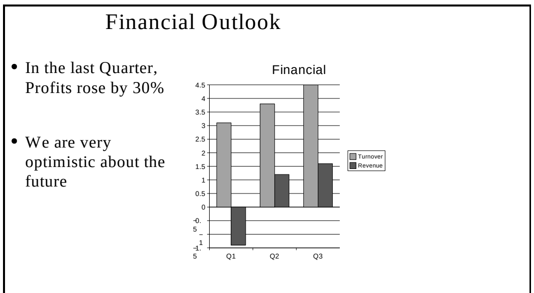

Figure 2: Common public version of slide 7

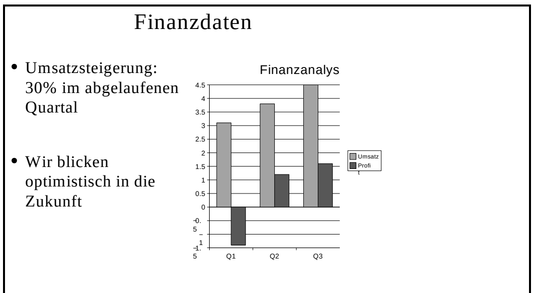

Figure 3: Personalised version of slide 7

ternative German version is sent to the people who asked for it. When the presentation reaches slide 7, for example, the presentation context manager selects two different versions for the public screen display (Figure 2) and the German-readers (Figure 3).

Tim uses his PDA to control the presentation. The Presentation Controller application shows him the current slide in reduced format, the thumbnails to give him context of where he is in his presentation, and the speakers notes which are for his eyes only, to prompt his voice.

During his talk, his PDA interrupts him with a video phone call. What could be important enough for this? Normally the fact that he is attending a seminar, let alone presenting the talk, would be enough for his Presentation Controller's awareness of the context to suppress the immediate delivery of any messages. But there are some higher priorities. . . "*Tim! your wife has had your baby, 6lb 3oz, 10.15am–all's well*".

3 Related Work

3.1 Content Personalisation

Software and content personalisation and internationalisation is now state of the art and most work has been based around the world wide web or other individual forms of multimedia presentation.

In an effort to combine wearable computing with context-driven applications, the Conference Assistant (Dey, Salber, Abowd & Futakawa 1999) is a prototype, for an extended, wearable PDA that assists people attending a conference. It provides various user interfaces for note taking, retrieving schedules and other conference material, based on location, content, and related context information.

The ADMITS project (Böszörményi, Hellwagner, Kosch, Libsie & Podlipnig 2003) is an experimental distributed system for the investigation of adaptive multimedia content. The system uses an adaptive query and presentation interface to the meta- and media database to present content based on MPEG-21 metadata (JTC1/SC29/WG11 2002, Vetro & Devillers 2002). These and similar approaches (Timmerer 2003, Steiger, Ebrahimi & Sanjuan 2003) use context information about user preferences to select the most suitable multimedia stream. Based on the profile of the viewing device, the delivered content can also be converted on the fly into the most suitable format matching the capabilities of the device.

The MPEG working groups (MPE 2001*a*, MPE 2001*b*) have also reached the stage of describing user preferences including choice of language, and the device capabilities for graphic, video and audio display (MPE 2002). Other frameworks for device description have been defined (Wireless Application Group 1999) or are still in development (Kly 2003). The CC/PP description, for example, has an example for print and display devices (from Appendix C in(Kly 2003)):

charWidth	integer	For a text display device the width of the character display. . . in typographical ens.
charHeight	integer	For a text display device the number of lines of text that can be displayed.
charset	Bag	Standard identifier for a set of characters such as UTF-8 or US-ASCII
pix-x	integer	For an image display device, the number of horizontal pixels that can be displayed.

pix-y	integer	For an image display device, the number of vertical pixels that can be displayed.
color	string	possible values are “binary”, “grey”, “limited”, “mapped” and “full”

MPEG has an essentially similar model of the display screen being defined by what is commonly called its “resolution”, i.e. x- and y- sizes in pixels.

Some attempts have been made to also automate the process of content generation, selection, or modification based on user preferences. Andrew Csinger, Kellogg Booth, and David Poole have outlined a process of structuring information so that selective content can be presented to different users, based on specific user models (Csinger, Booth & Poole 1994). Their approach is mainly aimed at a formalised process that allows the automated presentation of different content based on the user model and a semantic analysis of the available content.

Content adaptation for web pages has received the greatest attention to date. Most web servers today support serving web pages from the same URL in different languages. The decision about which languages are preferred for a web page is based on information passed by the web browser in the HTTP protocol header (Holtman & Mutz 1998, Fielding, Gettys, Mogul, Frystyk, Masinter, Leach & Berners-Lee 1999). Also state of the art is the possibility to design web pages that can adapt to different output devices using cascading style sheets (van Ossenbruggen, Hardman, Rutledge & Eliens 1997, Bos, Lie, Lilley & (Editors) 1998). More sophisticated approaches allow content to be structured and linked according to user preferences (Brusilovsky 1996) and habits (Armstrong, Freitag, Joachims & Mitchell 1995).

3.2 Identification and Positioning

The utility of a personalised presentation system depends on the ability of the system to identify the groups of people attending the presentation. There are a number of different techniques that allow the identification and localisation of people. The most ubiquitous location system, of course, is the Global Positioning System (GPS), which computes distance based on travel time differences between signals received from satellites orbiting the earth (Caplan 1996). GPS typically fails in indoor usage because it needs to maintain direct line of sight contact with the satellites

The Active Badge project (Want, Hopper, Falco & Gibbons 1992) envisioned a system for locating people in a work environment by letting them wear badges that transmit signals identifying the person in questions. Through collection points, the location information would be fed into a central computer. This central computer would then have an accurate representation of where employees were at all times. A similar location system was developed in the late 1990’s (Ward, Jones & Hopper 1997). Instead of active badges, this system uses ultrasonic sound transceivers to detect passive tags.

One of the biggest obstacles of these approaches was the necessity to wear a piece of hardware solely for identification purposes. This can be overcome if the technology that is carried around primarily fulfils a different objective, providing a means for positioning only as an additional feature. Considerable effort has, for example, gone into cellular telecommunication network based location or positioning systems. However, much of the location information, such as the cell of origin, signal strength, direction and time of signal arrival, and other triangulation data, is tightly held within the telecommunication infrastructure, primarily due to privacy concerns.

Recently, considerable research efforts have gone into local area positioning by the use of triangulation from multiple wireless network access points (Bahl & Padmanabhan 2000). While the volatility of the received signal strength limits the precision of these approaches, some successful attempts have been made at modelling these variations to achieve sufficient accuracy (Small, Smailagic & Siewiorek 2000, Ladd, Bekris, Marceau, Rudys, Kavraki & Wallach 2002).

4 Architecture

The design of the personalised presentation system is presented as a UML-like collaboration and synchronisation view in Figure 4. Messages between entities are numbered to show the sequence of dependencies.

There are many possible levels and dimensions of personalisation. It is particularly useful to distinguish two: content personalisation and presentation personalisation. We will briefly discuss the way that our system deals with both of these.

Firstly, and most simply is the matter of presentation. This applies where essentially the same content is to be made available to all users. However, the system can exploit different presentation modalities and forms to match the needs of the individual. This form of personalisation is already quite widespread on the web. For example, the typical web browser allows the user to define various aspects of the presentation of information, such as the font size, whether images will be displayed and the like. A particularly important form of such personalisation is critical for web accessibility for people with special needs, especially visual impairment. In the case of our system, some automated personalisation of this type has potentially high payoff. For example, suppose that some of the people at a presentation have low vision, so that they would be able to read the display if it has high contrast and large letters. In this case, our system should offer the presenter an easy means to automatically adapt the presentation, possibly breaking some ‘slides’ into two as necessary. This clearly introduces some serious user interface issues for the person creating a personalised presentation: they need to be able to control the degree of flexibility in presentation. It would clearly be quite undesirable for the a presenter to be surprised at the automated and uncontrolled changes to the public presentation they created.

Another important and widespread form of presentation personalisation involves translation. So, for example, if the presenter and most of the people at a personalised presentation speak English but just a few people in the audience find English difficult to understand, our system can deliver a translated form of the presentation to their personal device.

As soon as our system exploits the availability of personal devices, another form of personalised presentation will generally be needed. This follows from the different capabilities of the individual devices from powerful personal computers with a large screen to small personal organisers.

Although personalised presentation on personal devices involves elements of existing technology for rendering information and translation, there are still quite serious social and technical challenges. In terms of the personalisation architecture, we need to choose where the personalisation is performed. Some or all of it might be performed at the device used by the presenter. Equally, the presenter’s system could simply broadcast identical information to all the personal devices in the room and the personalisation could be performed locally by that device. The important difference between these two is that the former requires that the individual allow personal information to leave their machine so that it is available to the room and the presenter’s machine.

The other main type of personalisation is content personalisation. It means that different information will be available in the public or individualised presentation. So,

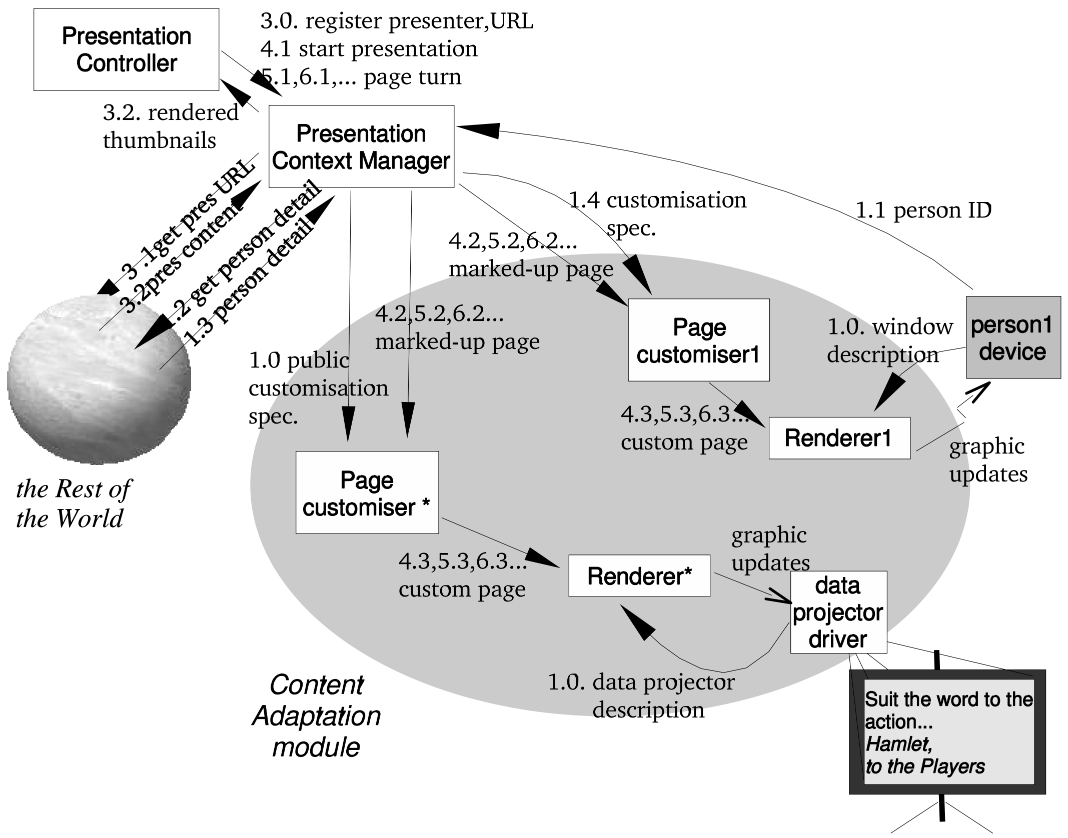

Figure 4: System components and their collaboration

for example, if one audience attends a personalised presentation, they will see one set of information while, at another time, a different group will see a different set. This has all the same issues associated with personal data. In addition, it introduces a number of quite challenging user interface and technical issues.

For the intended implementation of our system, the presentation content is created offline using typical presentation tools such as PowerPoint. In our approach we chose Impress, the presentation tool included with OpenOffice and StarOffice. The main reason for this is that PowerPoint uses a binary data format for presentations that cannot be easily modified or adapted. Since Impress contains a highly compatible import module, PowerPoint presentations can still be used in this system. Impress presentations are saved in an Extensible Markup Language (XML) format which makes them ideal for post-processing.

While the process of creating a presentation in principle remains the same as ever, individual pages can now be amended by tags or markup that specifies

- subtitles in different languages,
- subtitles for hearing impaired,
- the presentation notes for teleprompters, the presenter PDA, or live translators,
- what language a particular page is in,
- the security level for each page,
- a list of persons or Personas certain pages are restricted to or unsuitable for; and
- other personalised information such as subtitles or background information for specific individuals or groups.

In addition to generating such presentations manually in one or more steps, it is also feasible to enrich the presentation offline using automated tools, such as translation engines to create multilingual slides, database engines to retrieve background information, or similar tools.

Once the content has been created, it can be loaded into the presentation manager and the presentation can begin. The presentation manager is an application that runs behind the scenes in a local Intelligent Environment, a computing system “in the walls”. The presentation manager consists of two modules, a Presentation Context Manager, and a Content Adaptation module.

The Presentation Context Manager registers the presence of the presenter and the audience. This can be done through any means the local Intelligent Environment provides. What is important here is that it is not a requirement to know exactly which individual is currently present. This can be significant, because not everyone attending a presentation will necessarily agree to have their identity registered. From the viewpoint of the Presentation Context Manager, only the Persona needs to be identified. A Persona is a prototype of a group of persons, such as *the presenter*, *a person that can be trusted with confidential information*, *a person that prefers to receive their presentation in a foreign language*, and the like (Kay et al. 2002, Kay et al. 2003). The individual profiles and preferences of the Personas that are present determine the personalised presentations that need to be made available.

Depending on the information gathered by the Presentation Context Manager, one or more personalised presentations will be generated in the Content Adaptation module. The main presentation that will show up on the big screen is the smallest common denominator of what is preferred and acceptable for the whole audience. Content that is not suitable to be viewed by some members of the audience will be skipped or alternate versions will be displayed. An example for this are confidential data such as revenue figures within a company. Unless all participants are allowed to view this information, this part of the presentation will not be shown or it will be presented in a censored format prepared during content creation. Other examples include captions or foreign subtitles. If the majority of the audience requires Japanese subtitles and such subtitles are available for the presentation, they will be displayed on the main screen. Please note that the criterion that the majority of the audience decides the language of the subtitles is just an example. The actual decision depends on the preferences that have been set for a specific presentation or (if such preferences are missing for the presentation) the defaults that have been set for the presentation manager.

The big screen presentation is the minimum configuration for the content adaptation mechanism. A presentation with such a minimum configuration resembles a traditional presentation with the exception that the original content is modified, based on the profile of the audience. If that profile is known in advance, it would be possible for the presenter to modify her presentation accordingly (and indeed, this is what happens in many presentations now). In addition to this simple scenario, the content adaptation mechanism supports multiple output streams. The presenter, for example, could be equipped with a PDA that registers as an additional output device with the Intelligent Environment. This PDA could then, for example, be used as a teleprompter, allowing the presenter to prepare his speech as part of the presentation.

Similarly, it is possible for other output devices to register with the system as additional output channels. Such channels could either be generic channels or individual channels. Generic channels are channels that can be used by anybody participating in the presentation. Examples are audio streams from text to speech converters or simultaneous translators that are available via headphone, and live teleconference or web broadcasts of the current presentation. Additionally, it is possible for individuals in the audience to allow their PDAs to register with the presentation system. Based on these individuals' profiles and preferences, a personalised presentation can be transmitted to their devices. Examples for such personalised presentations are foreign language translations or subtitles that are shown on the PDA. Other examples include confidential content that may show up, for instance, on the personal device of the Chief Financial Officer of a company, but not on the main screen that is viewed by the rest of the audience. Figure 5 illustrates this for slide 7 from the scenario in Section 2.

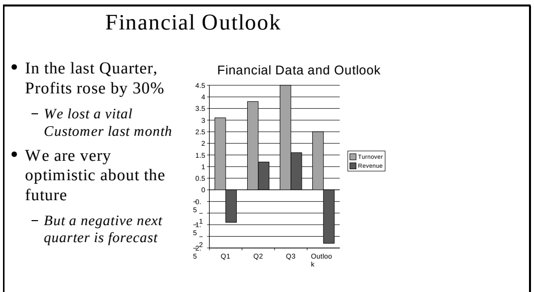

Figure 5: CFO version of slide 7

Finally, the mechanisms presented in this architecture are not restricted to output channels associated with persons, groups, or Personas. The same mechanism that provides cues on teleprompters for humans (e.g., the speaker), can also be used within the Intelligent Environment itself. Meta information about the state a presentation is currently in can, for example, be utilised to trigger lighting in a presentation theatre, cameras and multiplexers for teleconferences, and similar devices, allowing a presentation to make use of context information about the available infrastructure without requiring manual adaptation for every different venue it is used in.

5 Discussion and Conclusions

Our design of a system for personalising public presentations shows that current technologies are in theory capable of using context information of who constitutes "the audience" and their personal language and display rendering preferences to provide a personalised interface to a shared presentation.

The combination of technologies and standards has only become available very recently, and requires an infrastructure capable of handling context information to make it generally feasible. Implementation experience is needed before we can conclude whether these technologies are in fact sufficient for the task. Some concerns have already arisen.

The use of the current generation of display devices' capabilities contains hidden assumptions about what a "resolution" of a display implies. Where display resolutions are stated in pixels width and height, the relevant quantity for the readability of the display also depends on the distance from the display to the viewer's eye. With the wide range of distances that may apply in this situation (from the public projection display at tens of metres, down through individual's laptops at full armslength (700mm) or PDA at close armslength (450 mm)) and of screen "resolutions" in pixels, varying by a factor of two or more for the same physical screen size, the current capabilities are inadequate for satisfactory personalisation without some personal intervention by the user–the operation of selecting larger or smaller font for the same content and application on different monitors, familiar to many of us with middle-aged eyes losing some of their ability to accommodate even with the aid of correcting lenses.

Our model of personalisation and device customisation as a two-stage process working on content and then on rendering is sufficient for most presentation material, because there are typically (and ideally) few lines per slide (page) and few words per line. To personalise more densely packed pages raises the need to change the formatting significantly when rendering. Rendering may require more than pixel mapping: deciding what material to omit for the very small display, for example, has caused many problems with designing and delivering systems of Web page viewers for mobile phones, for example, and even at the scale of PDA displays, going from double column to single column. To perform well in this wider range of devices may require some of the device characteristics to be used at the earlier stage in the process as well.

The system we have described depends on trust for the customisation of content. Without some stronger security the system is inadequate for the presenter to know who is actually in the audience, since they may have faked their affiliations and be an invisible participant by wireless, concealed in the next room.

The concept of the persona is sufficient to protect personal privacy: it must be balanced with an equal amount of security and trust.

No trust is required for the display device formats. The customisation is potentially vulnerable to a form of denial of service attack: if a viewer registers a preference for a format that is extremely computationally expensive to supply, there may be some reduction in the speed of responding to page turning requests for all viewers, because of CPU cycles being unfairly taken up at the common server running the Content Adaptors. This seems a fairly remote possibility that is of little concern. It can be easily countered by setting the process priorities to render the common display image first, followed by the personal views.

Extensions of this scenario would add more input and output devices and channels. The context layer architecture is sufficiently generic to support multiple simultaneous input and output channels. Additional sensor information can add to the richness of personalised presentations.

An additional wall-mounted camera with a face detector and counter could further assist the presenter or session chair, for example, by allowing the Presentation Manager to determine the number of faces apparently present in the room. It could then present information about the discrepancies in identifying any anonymous attendees to the Presentation Controller.

Adding further sensors like this to the context is easy: deciding how to use them intelligently is as difficult as creating flexible, adaptive, humanly intelligent, context aware applications will always be. Once this architecture system is implemented many challenges in integrating it with a wider ranging intelligent environment and applications remain to be tackled.

Acknowledgements

The authors would like to acknowledge the support of the Smart Internet Technology CRC.

References

Armstrong, R., Freitag, D., Joachims, T. & Mitchell, T. (1995), Webwatcher: A learning apprentice for the world wide web, *in* 'AAAI Spring Symposium on Information Gathering', pp. 6–12.

Bahl, P. & Padmanabhan, V. N. (2000), RADAR: An in-building RF-based user location and tracking system, *in* 'Proceedings of the IEEE Infocom 2000', Vol. 2, Tel Aviv, Israel, pp. 775–784.

Bos, B., Lie, H., Lilley, C. & (Editors), I. J. (1998), Cascading style sheets, level 2, CSS2 specification, W3C Recommendation REC-CSS2-19980512, The World Wide Web Consortium.

Böszörményi, L., Hellwagner, H., Kosch, H., Libsie, M. & Podlipnig, S. (2003), 'Metadata driven adaptation in the ADMITS project', *EURASIP Signal Procesing: Image Communication, Special Issue on Multimedia Adaptation* .

Brusilovsky, P. (1996), 'Methods and techniques of adaptive hypermedia', *User Modeling and User-Adapted Interaction* **6**(2-3), 87–129.

Caplan, E. D. (1996), *Understanding GPS: Principles and Applications*, Artech House, Boston.

Csinger, A., Booth, K. S. & Poole, D. (1994), 'AI meets authoring: User models for intelligent multimedia', *Artificial Intelligence Review* **8**(5–6), 47–468.

Dey, A. & Abowd, G. (1999), Towards a better understanding of context and context-awareness, Technical Report GIT-GVU-99-22, Georgia Institute of Technology, Georgia.

Dey, A. K., Salber, D., Abowd, G. D. & Futakawa, M. (1999), The conference assistant: Combining context-awareness with wearable computing, *in* 'Proceedings of the Third International Symposium on Wearable Computers (ISWC 1999)', San Francisco, pp. 21–28.

Fielding, R., Gettys, J., Mogul, J., Frystyk, H., Masinter, L., Leach, P. & Berners-Lee, T. (1999), Hypertext transfer protocol – HTTP/1.1, Internet Society Standard RFC2616, The Internet Society, Network Working Group.

Holtman, K. & Mutz, A. (1998), Transparent content negotiation in HTTP, Internet Society Standard RFC2295, The Internet Society, Network Working Group.

JTC1/SC29/WG11, I. (2002), MPEG-21 digital item adaptation WD (version 2.0), Technical Report Document N4944, Motion Picture Expert Group.

Kay, J., Kummerfeld, B. & Lauder, P. (2002), Personis: a server for user models, *in* 'Proceedings of Adaptive Hypertext 2002', Springer-Verlag, pp. 203–212.

Kay, J., Kummerfeld, B. & Lauder, P. (2003), Managing private user models and shared personas, *in* K. Cheverst, B. N. de Carolis & A. Kruger, eds, 'UM03 Workshop on User Modeling for Ubiquitous Computing', pp. 1–11.

Kly (2003), Composite capability/preference profiles (cc/pp): Structure and vocabularies, working draft, W3C. http://www.w3.org/TR/2003/WD-CCPP-struct-vocab-20030325/.

Kummerfeld, B., Quigley, A., Johnson, C. & Hexel, R. (2003), Merino: Towards an intelligent environment architecture for multi-granularity context description, *in* 'Workshop on User Modeling for Ubiquitous Computing'. in press.

Ladd, A., Bekris, K., Marceau, G., Rudys, A., Kavraki, L. & Wallach, D. (2002), Using wireless ethernet for localization, *in* 'Proceedings of the 2002 IEEE/RJS International Conference on Intelligent Robots and Systems', IEEE Press, Lausanne, Switzerland.

MPE (2001*a*), MPEG-21 overview, Technical Report N4318, ISO/IEC JTC1/SC29/WG11, Sydney. http://ipsi.fraunhofer.de/delite/Projects/MPEG7/Documents/mpeg21-Overview4318.htm.

MPE (2001*b*), Overview of the MPEG-7 standard (version 5.0), Technical Report N4031, ISO/IEC JTC1/SC29/WG11. http://ipsi.fraunhofer.de/delite/Projects/MPEG7/Documents/w4031mpeg7overview.htm.

MPE (2002), MPEG-21 part 7: Digital item adaptation, draft N5353, ISO/IEC JTC1/SC29/WG11, Awaji Island, Japan. http://ipsi.fraunhofer.de/delite/Projects/MPEG7/Documents/mpeg21-Overview4318.htm.

Small, J., Smailagic, A. & Siewiorek, D. (2000), Determining user location for context aware computing through the use of a wireless LAN infrastructure, Technical report, Project Aura, Carnegie-Mellon University.

Steiger, O., Ebrahimi, T. & Sanjuan, D. M. (2003), MPEG-based personalised content delivery, *in* 'Proc IEEE International Conference on Image Processing'. to appear. Also at http://ltswww.epfl.ch/ steiger/publications/uma_icip.pdf.

Timmerer, C. (2003), Resource adaptation using XML within the MPEG-21 multimedia framework, Master's thesis, University Klagenfurt, Austria. http://www.itec.uni-klu.ac.at/ timse/stud/da.html.

van Ossenbruggen, J., Hardman, L., Rutledge, L. & Eliens, A. (1997), Style sheet support for hypermedia documents, *in* 'UK Conference on Hypertext', pp. 216–217.

Vetro, A. & Devillers, S. (2002), Delivery context in MPEG-21, *in* 'W3C Delivery Context Workshop', W3C World Wide Web Consortium, INRIA Sophia-Antipolis. http://www.w3.org/2002/02/DIWS/final.html.

Want, R., Hopper, A., Falco, V. & Gibbons, J. (1992), 'The active badge location system', *ACM Transactions on Information Systems* **10**(1), 91–102.

Ward, A., Jones, A. & Hopper, A. (1997), 'A new location technique for the active office', *IEEE Personal Communications* **4**(5), 42–47.

Wireless Application Group (1999), User agent profile specification, Technical report, Wireless Application Protocol Forum. http://www1.wapforum.org/tech/terms.asp?doc=SPEC-UAProf-19991110.pdf.