User Modeling and User-Adapted Interaction (2005) 00:1–42
DOI 10.1007/s11257-005-2304-5

Ontology-Based User Modeling in an Augmented Audio Reality System for Museums[*]

MAREK HATALA and RON WAKKARY
School of Interactive Arts and Technology, Simon Fraser University, Surrey, BC, Canada V3T 2W1. e-mail: {mhatala, rwakkary}@sfu.ca

(Received: 16 November 2004; accepted in revised form 13 August 2005)

Abstract. Ubiquitous computing is a challenging area that allows us to further our understanding and techniques of context-aware and adaptive systems. Among the challenges is the general problem of capturing the larger context in interaction from the perspective of user modeling and human–computer interaction (HCI). The imperative to address this issue is great considering the emergence of ubiquitous and mobile computing environments. This paper provides an account of our addressing the specific problem of supporting functionality as well as the experience design issues related to museum visits through user modeling in combination with an audio augmented reality and tangible user interface system. This paper details our deployment and evaluation of ec(h)o – an augmented audio reality system for museums. We explore the possibility of supporting a context-aware adaptive system by linking environment, interaction object and users at an abstract semantic level instead of at the content level. From the user modeling perspective ec(h)o is a knowledge-based recommender system. In this paper we present our findings from user testing and how our approach works well with an audio and tangible user interface within a ubiquitous computing system. We conclude by showing where further research is needed.

Key words. audio augmented reality, context-aware, museum guide, ontologies, semantic technologies, tangible user interface, testing, ubiquitous computing, user evaluations user modeling

1. Introduction

Fundamental to human–computer interaction (HCI) is the design of interactive systems that support people's goals and respond to individual backgrounds. In ubiquitous computing it is equally important to consider the influence of context on people's interactions and experiences. The intent is, as Fischer argues "to say the 'right' thing at the 'right' time in the 'right' way" (Fischer, 2001). A critical factor in ubiquitous computing is that what is perceived as "right" is largely mediated by the context within which the users find themselves.

In the area of user–adapted interaction, user modeling has attempted to address many issues related to HCI. Fischer provides a clear account of the successes and

---

[*] This paper has not been submitted elsewhere in identical or similar form, nor will it be during the first three months after its submission to UMUAI. (Submitted November 16, 2004)

future challenges of user modeling in HCI (Fischer, 2001). Among these challenges is the general problem of capturing the larger context in interaction (see Fischer, 2001, pp. 80–81). The imperative to address this issue is great considering the emergence of ubiquitous and mobile computing environments. This paper provides an account of addressing the specific problem of supporting functionality as well as the experience design issues related to museum visits through user modeling in combination with an audio augmented reality and tangible user interface system. We developed and tested a museum guide prototype, known as *ec(h)o* in order to research interaction design, user modeling, and adaptive information retrieval approaches that respond to the richness of a museum visit and the museum context.

Our aim is to support the limited input common to tangible user interfaces while maintaining rich and adaptive information output via a three-dimensional audio display. We believe an integrated modeling technique that is weighted toward modeling of implicit communication works well with a tangible user interface in creating a playful and discovery-rich experience. We believe this approach combined with ontologies and a rule-based system for information retrieval provides a richness of information that is responsive to the context and unique aspects of the museum visitor's interaction.

Our findings are both encouraging and cautionary. First, we found that it is possible to build a highly flexible and accurate user model and recommender system built on information collected from user interaction. This approach supported a user experience of liminal play and engagement. The ontologies and rule-based approach proved to be a strong combination. However, the ontological approach did not provide a clear enough contextual links between the artifacts and audio information and either more extensive knowledge engineering is needed or our approach has to be combined with stronger narration or discourse models.

In this paper we first review the general problem of context, our intended approach, and provide theoretical and related research as background. Following that we provide an account of our design and rationale for the prototype and its implementation. We give a detailed report of our evaluation and findings. We conclude with a brief analysis of our findings and discussion of future issues and research direction.

2. The Challenge of Capturing the Larger Context

Many HCI theorists and researchers identify issues of "context" as putting a strain on the traditional theories of HCI (Bodker, 1990; Dourish, 2004; Gay and Hembrooke, 2004; Nardi, 1995). As Nardi puts it, "we are beginning to feel a theoretical pinch, however – a sense that cognitive science is too restrictive a paradigm for finding out what we would like to know" (Nardi, 1995, p. 13).

For example, a visit to a museum reveals an everyday yet complex interaction situation. The factors within museum experiences are social, cultural, historical, and psychological. The influences on the experience vary from the actions and previous knowledge of the visitor, visitor's learning style, and the dynamics of others around them including friends, family and strangers. Naturally, the experience is also affected by the presence of the artifacts and collections, which are products of institutional history, curatorship, exhibition design, and architecture. The time of day, duration of visit, room temperature and so on all have an impact. The experience can be characterized as *multivariate*, that is, it cannot be assessed by a single factor such as exhibit design, signage, or time spent in front of an artifact (vom Lehn, et al., 2001). Instead, the museum experience is subject to multiple influences and results in multiple outcomes (Leinhardt and Crowley, 1998). Many similar situations have been discussed in design research such as how we work (Ehn, 1989), seek information (Nardi and O'Day, 1999), learn (Gay and Hembrooke, 2004), and live in our homes (Bell and Kaye, 2002; Tolmie et al., 2002).

In response to the issue of context, ethnographic and scenario-driven methods have begun to take hold in HCI practice (Carroll, 2000, 2002; Suchman, 1987). An emerging set of "context-based" theories for HCI has adapted ideas from an even wider spectrum of psychological, social, political and philosophical theories based on understanding human activity. For example, Nardi, Bødker, Gay and others (Bodker, 1990; Gay and Hembrooke, 2004; Nardi, 1995) have advocated on behalf of activity theory[1] . Dourish (2001, 2004) argues in his concept of embodied interaction that activity and context are dynamically linked – or "mutually constituent" (Dourish, 2004, p.14).

Suchman (1987) argues that the nature of interaction between systems and people require the same richly interpretive work required in human interaction, yet with fundamentally different available resources. For example, humans make use of non-verbal and inferential resources that can handle ambiguity and result in intelligible actions. This is not the case for computers. Fischer argues this raises two challenges: "(1) How can we capture the larger (often unarticulated) context of what users are doing (especially beyond the direct interaction with the computer system)? (2) How can we increase the 'richness of resources' available for computer programs attempting user modeling to understand (what they are told about their users) and to infer from what they are observing their users doing (inside the computational environment and outside)" (Fischer, 2001). In addition, Fischer cites Weiser and Bobrow (Bobrow, 1991; Weiser, 1993) in arguing that ubiquitous computing (and ultimately tangible user interfaces) aims to address the context issue by eliminating the separation between computational artifacts and physical objects, thus creating computational environments that require new approaches to interface and display.

---

[1] A theory developed by psychologists in the early 1920s (Vygotsky, 1925/1982), as a research tool and an alternative framework for understanding human activity as it relates to individual consciousness.

3. Background and Related Research

This research ties together several distinct domains that we will briefly review. These include adaptive museum guides, non-graphical user interfaces, user modeling, and semantic technologies.

3.1. ADAPTIVE MUSEUM GUIDE SYSTEMS

It is difficult to directly compare ec(h)o with other museum systems since our approach employs a unique form of interaction. However, ec(h)o shares many characteristics with the adaptive systems of HyperAudio, HIPS and Hippie (Benelli et al., 1999; Oppermann and Specht, 2000; Petrelli et al., 2001). Similar to ec(h)o the systems respond to a user's location and explicit user input. HyperAudio uses a static user model set by a questionnaire completed by the visitor at start-up time. HIPS and Hippie infer the user model dynamically from the interaction but they treat user interests as static. All systems adapt content based on the user model, location and interaction history. There are however many key differences between ec(h)o and these systems. HyperAudio, HIPS and Hippie depend on a personal digital assistant (PDA) graphical user interface (GUI), for example Hippie's audio interface is dependant on the GUI in such instances as *earcons* (Oppermann and Specht, 2000). ec(h)o uses an audio display as the only delivery channel, and a tangible user interface for input. Another difference lies in how the system generates response: ec(h)o uses inference at the level of semantic descriptions of independent audio objects and exhibit. ec(h)o extends the work of the Alfaro et al. (2003) by building a rich model of the concepts represented by the audio objects while HyperAudio and HIPS use partly pre-configured annotated multimedia data (Not and Zancanaro, 2000), and Hippie uses a simpler domain model. The last key difference is that ec(h)o treats user interests as dynamic, we look to evolving interests as a measure of sustainable interaction.

A museum guide that is conceptually more closely related to ec(h)o is the LISTEN project (Eckel, 2001), it is the follow-up to the Hippie system (Goßmann and Specht, 2002). It provides a personalized immersive audio environment delivered through wireless headphones. The LISTEN system is driven by the directional location tracking of the museum visitors and delivers "three-dimensional sound emitted from virtual sound sources placed in the environment" (Terrenghi and Zimmermann, 2004). The sound sequences are pre-processed by curators and artists. They are selected for the visitor based on a user-specified type. ec(h)o's user model changes dynamically based on the interaction. Its approach to the style of audio delivery and interaction model are also different. However, it is difficult to thoroughly compare LISTEN with ec(h)o as comprehensive evaluation results have not been reported beyond preliminary findings (Terrenghi and Zimmermann, 2004).

3.2. NON-GRAPHICAL USER INTERFACES

Prior to the evolution of adaptive and user modeling approaches in museum guide systems, there has been a strong trajectory of use of the PDA graphical user interface. Typically, hypertext is combined with images, video and audio (Aoki et al., 2002; Aoki and Woodruff, 2000; Proctor and Tellis, 2003; Semper and Spasojevic, 2002). Aoki and Woodruff have argued that in electronic guidebooks, designers are challenged to find the balance between burdening the visitor with the functions of selection, information management and contextualization (Aoki and Woodruff, 2000). The PDA graphical user interface approach comes at a cognitive and experiential cost. It requires the full visual attention of the visitor such that it is a competing element with the physical environment rather than a valued addition to that environment. Aside from projects like LISTEN, museum systems have mostly maintained the PDA graphical user interface approach despite the shifts in other domains to other approaches that better address the experience design issues most prominent in social, cultural and leisure activities.

Non-visual and non-graphical user interfaces, particularly audio display interfaces have been shown to be effective in improving interaction and integration with existing physical contexts. For example, Brewster and Pirhonen (Brewster et al., 2003; Pirhonen et al., 2002) have explored the combination of gesture and audio display that allows for complicated interaction with mobile devices while people are in motion. The *Audio Aura* project (Mynatt et al., 1998) explores how to better connect human activity in the physical world with virtual information through use of audio display. Audio is seen as an immersive display that can enrich the physical world and human activity while being more integrated with the surrounding environment. In addition, audio tends to create interpretive space or *room for imagination* as many have claimed radio affords over television. Audio augmented reality systems combined with tangible user interfaces often create very playful and resonant interaction experiences (Hummels and Helm, 2004). In fact, the distinction between augmented reality and tangible user interfaces can be blurry indeed (Ishii and Ullmer, 1997).

Tangible user interfaces like no other user interface concept is inherently playful, imaginative and even poetic. In addition, the concept has *immediacy* due to its physicality. Ishii and Ullmer's notion of *coupling bits and atoms* was informed by earlier work in graspable interfaces (Fitzmaurice et al., 1995) and real-world interface props (Hinckley et al., 1994). ec(h)o's tangible user interface draws on this notion by coupling an everyday and graspable object, a wooden cube with digital navigation and information (Ishii and Ullmer, 1997). Ishii was inspired by the aesthetics and rich affordances of scientific instruments (Ishii and Ullmer, 1997) and the transparency of a well-worn ping-pong paddle (Ishii et al., 1999). Simple physical display devices and wooden puzzles at the natural history museum where we conducted ethnography sessions inspired us as well.

In 1992, Bishop's Marble Answering Machine (Crampton-Smith, 1995) was an early embodiment of the immediate and playful qualities of tangible user interfaces. The prototype uses marbles to represent messages on the machine. A person replays the message by picking up the marble and placing it in an indentation on the machine. Ishii's PingPongPlus (Ishii et al., 1999) explores the intertwining of athletic play with imaginative play. The ping-pong table becomes an interactive surface. The ball movement is tracked and projections on the table of water ripples, moving spots, and schools of fish among other images react wherever the ball hits the table. While ec(h)o is more constrained in its play, the everyday wooden cube provides entry to a qualitatively diverse experience of interaction.

Over the years, various frameworks and interaction models have been proposed to better define tangible user interfaces. Holmquist and others (Holmquist et al., 1999) proposed defining concepts of containers, tools, and tokens. Ullmer and Ishii (Ullmer and Ishii, 2001; Ullmer, 2002; Ullmer et al., 2005) proposed a framework known as the MCRit that highlighted the integration of representation and control in tangible user interfaces. Shaer and others have extended MCRit to propose their Token and Constraints (TAC) paradigm (Shaer, 2004). Most relevant to our approach is Fishkin's proposed taxonomy which is situated and contextual in its thinking (Fishkin, 2004). Fishkin's taxonomy is a two-dimensional space across the axes of *embodiment* and *metaphor*. Embodiment characterizes the degree to which "the state of computation" is perceived to be in or near the tangible object. Metaphor in this sense is the degree to which the system's response to a user's action is analogous to a real-world response to a similar action. Further, Fishkin divides metaphor into *noun metaphors*, referring to the shape of the object, and *verb metaphors*, referring to the motion of an object. For example, in ec(h)o, according to Fishkin's taxonomy embodiment would be considered "environmental" since the computational state would be perceived as surrounding the visitor given the three-dimensional audio display. In regard to metaphor, ec(h)o would be a "noun and verb" since the wooden cube is reminiscent of the wooden puzzle games in the museum and the motion of the cube determines the spatiality of the audio as turning left in the real-world would allow the person to hear on the left.

3.3. USER MODELING

'Knowledge-based HCI' (Fischer, 2001) explores the possibility of implicit communication channels between a human and a computer. These channels capture the idea of shared knowledge about problem domains, communication processes, and agents involved with communicating parties. This notion is very close to the goals of user modeling (Wahlster and Kobsa, 1989). Several researchers worked on the incorporation of user modeling in order to improve the collaborative nature of human–computer systems (for examples see Fischer, 2001). In our research we expand the role of user modeling into the realms of audio augmented reality and tangible user interfaces.

In the context of our work, the user model performs the function of a recommender system (Resnick and Varian, 1997). "Recommender systems represent user preferences for the purpose of suggesting items to purchase or examine" (Burke, 2002). Several types of recommendation techniques have been developed: collaborative, content-based, demographic, utility-based, and knowledge-based. Often the researchers combine several techniques to achieve maximum effect. Burke (2002) compares the recommendation techniques from the perspective of their ability to deal with the 'ramp-up' problem (Konstan et al., 1998): an introduction of new users and new items. In this regard, knowledge-based recommenders perform favorably. This is an important feature for ubiquitous computing environments that often manifest the 'walk-up-and-use' characteristic. Knowledge recommender systems require three types of knowledge (Burke, 2002): catalog knowledge or knowledge about objects to be recommended, functional knowledge of mapping between user needs and objects, and user knowledge. In the case of ubiquitous computing applications the functional knowledge must include the knowledge of the environment since context-awareness is a key requirement of ubiquitous computing systems. The knowledge of the user can be specific to the domain of recommendation; or can expand to general user modeling.

From a user modeling perspective, ec(h)o is a knowledge-based recommender system. Similar to Towle and Quin's (2000) proposal, we build explicit models of users and explicit models of objects. However, in ec(h)o the models are not built around specific content but rather ec(h)o uses ontologies at a higher level of abstraction. Users, objects, and environment are annotated with these ontologies. Another significant feature where ec(h)o differs from other knowledge-based recommender systems (for example Entrée, Burke, 2002), is that it does not solicit user's feedback about the quality of recommendations.

In addition to user modeling, capturing user interests is a central research focus of several disciplines such as information retrieval and information filtering. Most such systems are based on document retrieval where a document's content is analyzed and explicit user feedback is solicited in order to learn or infer user interests. In our approach, there is no direct feedback from the user. Our prototype can be categorized as a personalized system, as it observes user's behavior and makes generalizations and predictions about the user based on their interactions (Fink and Kobsa, 2002; Seo and Zhang, 2000). Our approach to observation of user behavior is unobtrusive, similar to approaches to monitoring user browsing patterns (Lieberman, 1995; Mladenic, 1996) or user mouse movement and scrolling behavior (Goecks and Shavlik, 2000).

3.4. SEMANTIC TECHNOLOGIES

Modeling is an integral part of the user modeling by definition. Several types of models are used ranging from simple categories through statistical models, Bayesian networks to formal knowledge models as known in symbolic artificial intelligence (Wahlster and Kobsa, 1989). It is these latter models that potentially benefit the most from semantic web research.

The semantic web initiative (Berners-Lee et al., 2001) aims to achieve a vision of creating a web of meaning. It argues for a set of technologies and techniques that integrates artificial intelligence into the core of the World Wide Web. The cornerstone of semantic web is ontologies (Chandrasekaran, et al., 1991) that provide a mechanism for modeling domains of interest. The formalization is essential for reasoning (Post and Sage, 1990) about the domain. Ontologies and reasoning are basic semantic web technologies that are useful not only in traditional web application domains such as knowledge management, data integration and exchange, or agent coordination but are extensively used in other domains for representation purposes. For example, Baus and colleagues (2002) use ontologies to model the environment in a mobile navigation system. In the Story Fountain system (Mulholland et al., 2004), ontologies are used to describe stories and the domain in which they relate. In order to determine the appropriate domain, reasoning is employed for the selection and organization of resources from which the stories are built.

A main advantage of ontologies, as the concept has developed within semantic web research is the ability to cross-link different domains (Noy and Hafner, 1997). In the area of user interaction this provides us with a clear formalism to connect knowledge about the user, environment, and user aims.

An obstacle in connecting and sharing data, is that often the knowledge captured within an application is at too low a level of abstraction; it is too domain specific. Ontologies provide a mechanism for building several layers of abstraction into the model (Noy and Hafner, 1997).

The assumption we are testing in our approach is that we can use ontologies and semantic web techniques to build interactive systems that successfully operate at higher levels of abstraction. Such a design can be shared across multiple applications. Furthermore, only low-level application-specific logic has to be developed for a new application. Our approach tests this assumption in the context of an audio augmented reality system with a tangible user interface.

4. Design and Rationale

The aims of our design were to develop a ubiquitous computing museum guide that supports *liminal* and engaging play in its user experience; investigates user modeling limited by implicit input from users' actions; and delivers a wide breadth of information associated with artifacts on exhibit via audio display that is responsive to users' changing interests. In short, we aimed to investigate less explored avenues in current museum guide systems research including play, embodied interaction, and highly associative as well as contextualized content delivery.

In the last decade, advances in audio museum guides include visitor–driven interaction, access to large collections of supplementary information for museum artifacts, and the development of adaptive and context-aware systems. Many of these advances have come on the heels of innovations in mobile computing including computer processing capabilities, data storage, connectivity and size. This has culminated in the growing use of PDA devices combined with sensor systems for use as interactive museum guides (Proctor and Tellis, 2003). Yet, outside the domain of museums, for example in the area of games and ubiquitous computing, Björk and his colleagues have identified the need to develop past end-user devices such as mobile phones, personal digital assistants and game consoles (Bjork et al., 2002). They argue that we need to better understand how "computational services" augment games situated in real environments. Our design ethnography observations confirmed that museum interactives such as computer kiosks were less used than physical and play-based interactives (Wakkary and Evernden, 2005). In addition, Proctor (Proctor and Tellis, 2003) has found that in museum use PDAs create expectations of a multimedia experience that lessens the relationship between the visitor and the artifacts. As examples, visitors tend to want more of everything yet they quickly lose interest in audio/visual and interactive clips; the visual screen made the moments in-between interactions problematic since if the screen became blank, visitors thought the devices were broken, yet they did not want the screen on all the time since it distracted them from the exhibition. The main point of these findings is that the focus of the visitor is on the experience of the device rather than the experience of the museum.

The anthropologist Genevieve Bell has described museums in terms of *cultural ecologies* (Bell, 2002). Bell sees the museum visit as a ritual determined by space, people and design. She decomposes the visiting ritual into three observational categories: space, visitors, and interactions and rituals. Different types of museums have different ecologies, for example Bell describes different attributes in each of the observational categories between art museums and science museums. These ecologies are seen to be distinct and supportive of different kinds of museum visits. Bell also describes concepts that are common to all museum ecologies. We have drawn on and extended two of these concepts in developing our approach, *liminality* and *engagement*.

Liminality defines museums as places that embody an experience apart from everyday life. Positive museum experiences are transformative, spiritual, and even moving. A museum visitor should be inclined to pause and reflect, thus liminality can be seen to permit a deeper engagement. Engagement is a key concept for museums as people go to museums to learn, however this engagement is often packaged in an entertaining way; museums are a balance between learning and entertainment spaces. It is easy to see how liminality and engagement include ludic experiences in which play and discovery are encouraged. In our adult lives, play is an experience set apart from our everyday activities: Huizinga refers to play as invoking a "magic circle", a liminal space for games (Huizinga, 1964); Carse describes "deep play" as a profound level of ritualized engagement causing reflection on everyday experiences (Carse, 1987); and psychologist Csikszentmihalyi has described "flow" as a high level of engagement, risk and challenge found in play (Csikszentmihalyi, 1990).

Our aims led us to a design that was inherently minimal and playful. In order to move past the limitations of device-centered approach we developed a tangible interface supported by an audio display, and a user model and adaptive information retrieval system. The tangible interface creates a playful transition between the physical space and the virtual information space of the audio. The audio display creates a virtual context that allowed us to create new layers of engaging experiential spaces such as ambient sounds and conversational information delivery.

Given the limited input and output of our interface, we chose a user model approach to act as a mediator for the visitor. The user model dynamically integrates movement interaction and visitor content selection into initial pre-selected preferences. Based on this dynamic model we could infer potential interests and offer a corresponding range of content choices even as visitors' interests shifted over time. In addition, the use of semantic technologies allowed for coherent and context responsive information retrieval.

While arguably other interface approaches could have been utilized in conjunction with the integrated modeling technique, such as a simple push-button device for input or a mobile text display device for output, such a strategy would be incongruent with our experience design goals. Nevertheless, we designed our user modeling and semantic technologies technique such that it could be easily modified for other interfaces and applications.

The project was informed by ideas of ecologies, like Bell's *cultural ecologies* and prominently used audio. This combination led us to the name ec(h)o, which is intended to signify the words *eco*, an abbreviation for the word "ecology", and *echo*, denoting the acoustic aspects of the project.

4.1. VISITOR SCENARIO

In order for us to better describe the system we developed, we provide below a typical visitor scenario. It should be noted that the scenario describes aspects of the project that are not the focus of discussion in this paper such as soundscapes. The scenario refers to an exhibition about the history and practice of collecting natural history artifacts in Canada at the Canadian Museum of Nature in Ottawa:

> *Visitors to the Finders Keepers exhibition can use the ec(h)o system as an interactive guide to the exhibition. Visitors using ec(h)o begin by choosing three cards from a set of cards displayed on a table. Each card describes a concept of interest related to the exhibition. The cards include topics such as "aesthetics", "parasites", "scientific technique" and "diversity". A visitor chooses the cards "collecting things," "bigness," and "fauna biology." She gives the cards to an attendant who then gives the visitor a shaped wooden cube that has three colored sides, a rounded bottom for resting on her palm and a wrist leash so the cube can hang from her wrist without her holding it. She is also given a pair of headphones connected to a small, light pouch to be slung over her shoulder. The pouch contains a wireless receiver for audio and a digital tag for position tracking (see Figure 1).*
>
> *Our visitor moves through the exhibition space. Her movement creates her own dynamic soundscape of ambient sounds. As she passes a collection of animal bones she*

Figure 1. A Museum visitor testing the ech(o) system.

hears sounds that suggest the animal's habitat. The immersive ambient sounds provide an audio context for the collection of objects nearby.

As she comes closer to a display exhibiting several artifacts from an archaeological site of the Siglit people, the soundscape fades quietly and the visitor is presented with three audio prefaces in sequence. The first is heard on her left side in a female voice that is jokingly chastising: "Don't chew on that bone!" This is followed by a brief pause and then a second preface is heard in the center in a young male voice that excitedly exclaims: "Talk about a varied diet!" Lastly, a third preface is heard on her right side in a matter-of-fact young female voice: "First dump … then organize." The audio prefaces are like teasers that correspond to audio objects of greater informational depth.

The visitor chooses the audio preface on the left by holding up the wooden cube in her hand and rotating it to the left. This gesture selects and activates an audio object that is linked to the audio preface of the scolding voice warning against chewing on a bone. The corresponding audio object delivered in the same female voice yet in a relaxed tone, is about the degree of tool making on the part of the Siglit people: "Artifact #13 speaks to the active tool making. Here you can actually see the marks from the knives where the bone has been cut. Other indicators include chew marks … experts are generally able to distinguish between rodent chew marks and carnivore chew marks."

After listening to the audio object, the visitor is presented with a new and related audio preface on her left, and the same prefaces are heard again in the center and to her right. The audio prefaces and objects presented are selected by the system based on the visitor's movements in the exhibition space, previous audio objects selected, and her current topic preferences.

4.2. INTERACTION MODEL

Our interaction model relies on a turn-taking approach generally based on the structure of a conversation[2]. We designed our audio objects in two parts,

---

[2] We use the term "conversation" in the context of the use of conversation analysis to inform HCI design. The idea of using conversation analysis concepts as a structural metaphor for non-speech interfaces is not unique in HCI, see for example (Norman and Thomas, 1990).

prefaces and *audio objects*: *prefaces* act as multiple-choice indices for the more detailed *telling* of the audio object. The tangible user interface provides input for a response to the delivery of *prefaces*.

The implementation went as follows: *ec(h)o* offers the visitor three short audio pieces as *prefaces*. The system is in effect offering three turn-taking possibilities for the visitor. Switching between the stereo channels created localization: we used the left channel audio for the left, right channel audio for the right, and both channels for the center. It is a simple *egocentric* (Brewster et al., 2003) spatial structure that allows the three *prefaces* to be distinguishable and an underlying content categorization structure to exist. The spatialization was mapped to the tangible user interface for selection. The visitor *responds* by rotating the wooden cube in his hand and thus selecting a *preface*. The system delivers the audio object related to the *preface*. After the delivery of the object, the system again offers three *prefaces*. The visitor's response is expressed through the gesture selection with the wooden cube. Additionally, the system may be met by no response, because the visitor does not wish to engage the system. The system will then enter into a silent mode. The visitor may also have moved away and the system will then initiate a soundscape.

The *prefaces* were written to create a sense of surprise and discovery. The audio recordings used a diverse set of voices that were informal in tonality and style. This added to the conversational feel and created an imaginary scene of a virtual cocktail party of natural historians and scientists that followed you through the museum. The audio objects were developed through interviews with museum staff and researchers (Wakkary et al., 2004).

A topic of interest is conceptually represented by each preface or spatial location. The structure is very simple given the limited choices of three options. The navigation is as follows: a visitor is played three *prefaces,* one to his left, another to his center and the third to his right. He selects the *preface* on his right side and listens to the linked audio object. On the subsequent turn the visitor hears the same two *prefaces* he did not select, and again he hears them to his left and to his center. Since he previously chose the *preface* to his right he now hears a new *preface* in that location. If the visitor then selects the center *preface,* on the subsequent turn only that *preface* is replaced by a new *preface* in the center position. If a *preface* has been replayed three times without being selected, it is replaced by a *preface* linked to an audio object of a completely new topic.

The audio objects are semantically tagged to a range of topics. At the beginning of each interaction cycle, three audio objects are selected based on ranking using several criteria such as current levels of user interest, location, interaction history, etc (see Section 4.4.2). The topics of each object are not explicit to the visitor; rather the consistency and content logic are kept in the background.

In regard to the design process, many of the design choices were made through a series of participatory design workshops and scenarios, details of which have been written in another paper (Wakkary, 2005, in press). For example, the tangible user interface and its implementation as an asymmetrically shaped wooden cube resulted from these workshops. We also recreated the exhibition environment in our labs; this aided us in the design the interactive zones and audio display.

4.3. USER MODEL

At the core of the ec(h)o's reasoning module is a user model (Wahlster and Kobsa, 1989) that is continually updated as the user moves through the exhibition space and selects audio objects.

Figure 2 shows an interaction schema of the user model with other modules. There are two main update sources in the system. First, as the user moves through the exhibition the speed of the movement and/or stops in relation to different artifacts provides updates to the user model. The user type is computed based on the speed and uniformity of the user movement. The slowing down and rest points in front of an artifact are interpreted as an interest in concepts represented by the artifact.

The second source of updates to the user model considers a user's direct interaction when selecting an audio object. In the model this correlates to an increased interest on behalf of the visitor in concepts presented by the audio object and this is reflected in the user's interaction history.

4.3.1. *User Model Components*

Interaction history is a record of how the user interacts with the augmented museum environment. Two types of events are stored in the interaction history: the user's movement and user's selection of objects. The user path through the museum

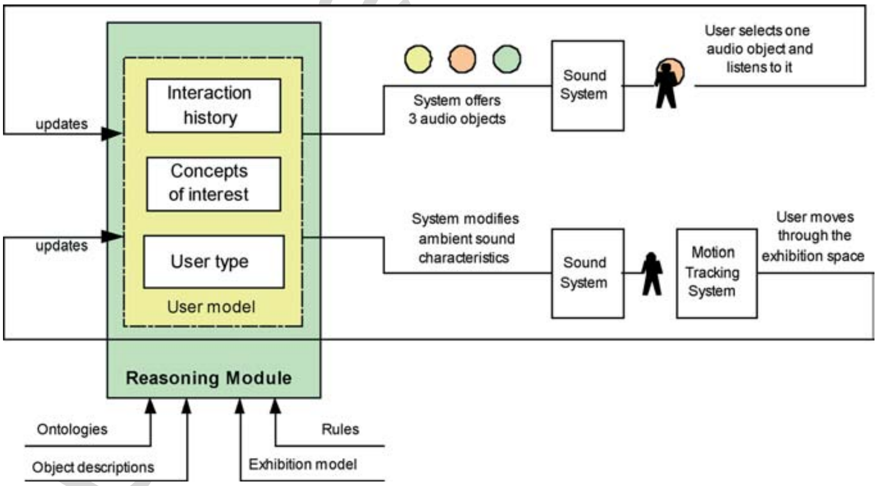

Figure 2. Interaction of usermodel with other modules.

is stored as discrete time-space points of locations on the path. A second type of information stored in interaction history is the user's selections of audio objects.

User type in the museum context is well studied in museum studies (Dean, 1994) and is used in several systems personalizing the user experience (Serrell, 1996; Sparacino, 2002). In the case of ec(h)o, several categorizations were used, for example one user may review almost every artifact on her path, and another user may be more selective and choose artifacts that have only certain concepts. Our categorization of user types is based on Sparacino's work (Sparacino, 2002). It classifies users into three main categories. These categories were validated by our site studies and interviews with staff at various museums:

- The avaricious visitor wants to know and see as much as possible. He is almost sequential, and does not rush;
- The selective visitor explores artifacts that represent certain concepts and is interested in only those concepts;
- The busy visitor does not want to spend much time on a single artifact preferring to stroll through the museum in order to get a general idea of the exhibition.

In ec(h)o, the user type category is not static but is updated every minute. The rules for the type specification consider the location data accumulated within the longer time interval and concepts of previously selected audio objects.

User interests are represented as a set of weighted concepts from the 'concept ontology'[3] (described in Section 4.4). In ec(h)o, each artifact and exhibition is annotated with a set of concepts from the same ontology. The audio objects present a set of particular concepts as well. In each interaction step the system updates the user interests in response to two update channels described above. The update process is described in detail in Section 5.5.

The interaction of the user with artifacts and audio objects is stored in the interaction history that together with the user types are used to infer the user's interests. Several aspects of the update process are parameterized. We discuss the user model parameters and the user model update process in Section 4.5 after we introduce the model for representing content and context in the next section.

4.4. INFERENCE-BASED AUDIO OBJECT RETRIEVAL

The audio object retrieval process is performed by the rules that encode multiple object selection criteria. The rules match semantic descriptions of the objects and the museum environment with user information maintained by the user model.

---

[3] We use term 'interest' or 'user's interest' when referring to the user model. We use the term 'concept(s) of interest' when referring to the concepts when used to annotate the objects or before they were used to modify the level of corresponding interests in the user model. The relation of the interest in the user model to the concepts in the concept ontology is crucial as it links user model to the model representing content and context as described in the subsequent section.

The content model is based on the semantic description of all the properties of the audio objects and the museum environment that could help us to select visitor and context relevant audio objects. Our ontological model builds significantly on the standard Conceptual Reference Model (CRM) for heritage content developed by CIDOC (Crofts et al., 2003). The CRM provides definitions and a formal structure for describing the implicit and explicit concepts and relationships used in the cultural heritage domain. We have also developed several ontologies specifically for the purpose of ec(h)o.

4.4.1. *Ontologies for Describing Content and Context*

The content of the audio object is not described directly but annotated with three entities: concepts of interest, topics, and themes. The *concepts of interest*[4] describe the domains that are expressed by the audio objects such as ‘evolution’, ‘behavior’, ‘lifestyle’, ‘diversity’, and ‘habitat’. We realized that it would be impossible to model the content at the actual descriptive level of objects, science and events, so we opted for higher levels of abstraction that in turn provide a unifying degree of formalization for all audio objects in the collection. The starting point for our concept ontology was a set of concepts used by the museum curators at the time of designing the exhibit. We have further extended this initial ontology with concepts identified through analysis of the content of audio objects used in ec(h)o and through interviews with museum researchers (Wakkary et al., 2004). As a result the concept ontology has a flat structure with 39 identified concepts[5]. These concepts are mapped to the Dewey Decimal Classification (represented as an ontology), which indirectly gives our concept ontology a hierarchical structure that can be used for drawing inferences.

The concepts play a significant role in the system in linking audio objects and museum artifacts with user interests. The user model (described in the section above) captures a level of user interest in each concept. The audio object retrieval mechanism uses those levels to determine the most appropriate audio objects for the next interaction turn. Similarly, the exhibits are annotated with the concepts that are visually represented in the exhibit (so called *visual concepts*). When a visitor slows down or stops in the exhibit those visual concepts are used to update the user model.

A *topic* is a higher-level category for describing several objects within the same exhibit. Objects annotated with different concepts of interest can still have the same topic. *Themes* are defined as entities that are represented across several exhibitions and are supported by one or more topics; for example, the theme of ‘bigness’ can include topics such as ‘invertebrates’ and ‘marine biology’.

---

[4] The concepts of interests represent interests as used by the user model introduced in Section 4.3.

[5] As a result the concrete user model can contain up to 39 interests. However, this is very unlikely as a result of the implemented user model update process described in Section 4.5.

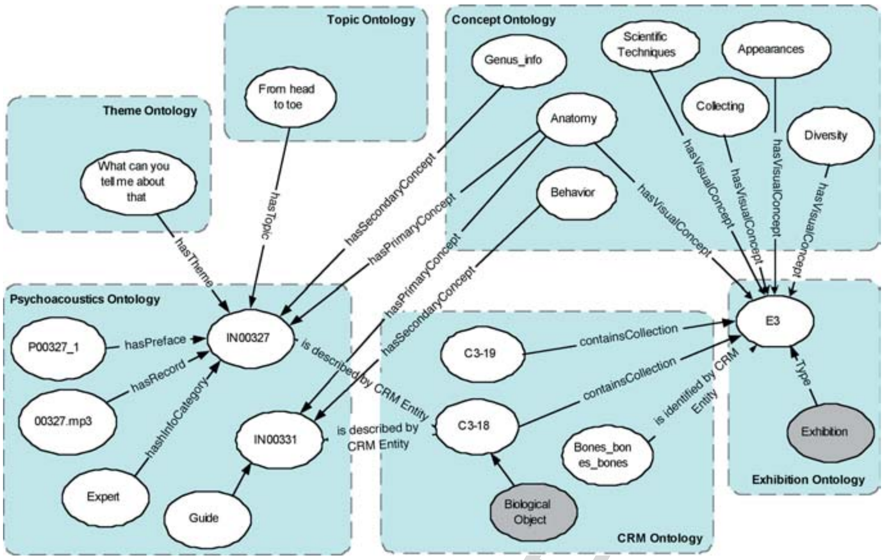

Figure 3. ec(h)o content ontologies.

We have used *CRM* to describe the museum exhibits and artifacts. CRM provides a comprehensive model for describing physical entities, temporal entities and places. We have used CRM to model events and places related to the objects and narratives captured in the audio objects.

Figure 3 shows an example how audio objects (‘IN00327’ and ‘IN00331’) are represented in ec(h)o. Both objects exist as independent entities and are related through several ontological relations. The audio object ‘IN00327’ is annotated with the concepts of interest ‘Anatomy’ and ‘Genus Info.’ ‘IN00327’ has a topic ‘From Head to Toe’ and supports the theme ‘What Can You Tell Me About That’. The audio object ‘IN00331’ is annotated with the concepts of interest ‘Anatomy’ and ‘Behavior’ but is a ‘Guide’ object (some relations for ‘IN00331’ were omitted from the picture). The ‘Guide’ objects differ from the ‘Expert’ objects by being directly related or referring directly to the artifacts in the exhibition, while the content of ‘Expert’ objects describes more general knowledge and is reusable in different contexts.

Both objects ‘IN00327’ and ‘IN00331’ describe the same museum artifact ‘C3-18’ representing a ‘common dolphin skull’ artifact in the exhibition ‘E3’. The ‘C3-18’ is an instance of a ‘Biological object’ class in the CRM and has many properties that link it to other artifacts in the exhibition (not shown in the picture). The exhibit instance ‘E3,’ from the exhibit ontology holds the information about the artifacts in the particular exhibit. In addition, ‘E3’ is annotated with visual concepts ‘Collecting’, ‘Anatomy’, ‘Scientific Techniques’, ‘Diversity’ and ‘Appearances’ that are represented visually in this particular exhibit.

Both topics and themes are common tools used by the curators when designing a museum exhibition. In ec(h)o, we use topics and themes in the audio object selection process to support fluency of the interaction between the user and the system. We use CRM referents of place and time period of the artifacts for the selection of the corresponding background sounds appropriate for the presented audio objects.

4.4.2. *The Audio Object Selection Process*

The audio object selection is based on the ranking of objects. Multiple criteria contribute to the ranking and the audio object with highest ranking is selected. The ranking criteria reflect the dynamic nature of the interaction that is represented by a level of current user interests, previously listened to audio objects and exhibits visited. The system is not intended to be a guide system but rather to enrich the experience of the exhibit and artifacts.

The ranking criteria are listed in Table I. Criterion 1 contributes to audio objects by further describing previously described artifacts while criterion 2 contributes to the ranking of guide audio objects if a previous audio object was also a guide audio object or the user entered a new exhibit[6]. Criteria 3–5 provide for the continuity in the interaction by contributing to the audio objects that elaborate on the same concepts within the same topic and theme. The contribution of criterion 6 is scaled with the current levels of user interests (which change after each interaction step).

The selection process is parameterized and the contribution of each criterion is weighted by its relative importance. Instead of doing extensive testing for weight values the weights were established in consultation with an expert in interactive narrative and storytelling. Table I shows the relative weight distribution for ranking criteria. The only criterion, which we have tested for a range of values, is the contribution between matching concepts of interest in the user model and matching audio object descriptions (Criterion 6, see Section 7 for evaluation and testing results). The remaining values were kept stable. The 'From', 'ec(h)o', and 'To' columns show the absolute values for the weights and '%' column show the relative contribution to the overall ranking[7]. The 'From' column shows the absolute values for the weights when interests in the user model contributed to the object ranking, at a minimum of 13% and 'To' column shows the weight values when interests contributed, up to a maximum of 48%.

---

[6] Guide objects provide for quick orientation in an exhibit with multiple artifacts by directly referring to those artifacts.

[7] The objects score in all criteria, otherwise the percentage contribution is shifted towards the matched criteria. Also, it should be noted that while criteria 1–5 always contribute their full weight the contribution from the criterion 6 varies. The value of criterion 6 shown in the table is the user level of interest in the audio object represented at the maximum level.

Table I. Weight distribution for object ranking

Criteria	From	%	ec(h)o	%	To	%
1. Describing artifact previously referred to by the audio object	10	22	10	16	10	13
2. Object is a 'guide' type of audio object describing an artifact	6	13	6	10	6	8
3. Continuing in previous topic	8	18	8	13	8	11
4. Continuing in previous theme	8	18	8	13	8	11
5. Continuing description of concepts in previous audio object	7	16	7	11	7	9
6. Concepts in the object match user interest	5.6	13	22	36	36	48

The middle column labeled 'ec(h)o' shows the actual values used in the final demonstration. The distribution of ranking contributions in the 'ec(h)o' column is used for audio object selection while a visitor remains within the same exhibit. When users change exhibits only the criteria 2, 3, 4, and 6 are used with the relative distribution of 14, 18, 18, and 50%[8] respectively.

The criteria are implemented in the form of forward chaining rules in which the condition part matches semantic characteristics of each audio object with the interaction history and user interests. If the characteristics of the audio object satisfy the condition, the rule is fired and the ranking for the object is increased. Several rules can be fired for the same audio object. After all rules for the matching audio objects are fired and contributed to ranking, the object with the highest ranking is selected.

For example, the rule below represents criterion 1 in Table I. The rule adds ratings to the audio object that describes the same artifact as the object being replaced. The rule checks whether candidate object `?in2` describes the same artifact `?a` as previous object `?in1`. Next, we make sure that `?in2` is not an exhibition object but an actual artifact within the exhibition. The `PropertyValue` is a fact representing semantic descriptions in the form of triples (obtained from the ontologies via transformation when loaded into the inference engine). For brevity, we have also used XML entity descriptions to refer to the namespaces of the ontologies.

---

[8] It should be noted that the levels of interests in the user model are updated with visual concepts in the new exhibit *before* they are used to calculate the ranking. As a result the influence of the context of the new exhibit (in addition to 14% for guide objects) is strongly represented in a 50% contribution from the user model.

```
(defrule artifact2artifact- - -1
  (user-group (user ?u) (group 1))
  (replace (user ?u) (context ?e) (object ?in1) (context ?e)
           (sequence ?seq) (time-chosen ?t))
  (test (neq ?in1 nil))
  (in-context ?a ?e)
  (PropertyValue &psch;#describes ?in1 ?a)
  (PropertyValue &psch;#describes ?in2 ?a)
  (not (PropertyValue &rdf;#type ?a &crm;#exhibition))
  (not (replaced (user ?u) (next-object ?in2) ))
=>
  (call ?*object-ratings* addRating ?u ?in1 ?in2 ?
           *artifact-rating* ?t))
```

For more details about representation and information retrieval aspects in ec(h)o see (Hatala et al., 2004) and (Hatala et al., 2005).

4.5. USER MODEL UPDATE PROCESS

The rule-based user model provides a generic structure that enables the system developer to consider several inputs that influence user interests. In addition, the model allows the developer to tune the relative influence of each input using a set of parameters. In ec(h)o, we interpreted two aspects of the user interaction with the system and environment: user movement and audio object selection. Each of these actions has a different effect on the model of user interests.

Influence of initial interest selection. A new user starts with a blank user model. In order to bootstrap the model we ask each visitor to indicate initial interests. Prior to entering the exhibition space the user selects a set of cards representing concepts of interest that best match their interests (see Section 4.1 Visitor Scenario). An operator enters the chosen concepts of interest into the user model as user's initial interests[9] and from that point the system evolves the user model through the two update channels described below. The parameter controlling the initial interests' weight can be set by the developer.

Influence of object selection on user interest. In ec(h)o each audio object is described by two concepts of interest: primary and secondary. When a user selects an audio object its primary and secondary concepts of interest are used to update the corresponding user's interests if they were already present in the user model, or adds them to the user model if they were not previously included in the model. As a result, the model is *dynamic* and the number of interests in the model can vary

---

[9] In a fully implemented system the same could be achieved automatically by asking the user to select a set of initial interests using a computer kiosk system.

depending on each user's individual interaction with the system. The model enables the developer to specify the parameters of how much the primary and secondary concepts of interest in the selected audio object increase the level of corresponding interests in the user model.

Influence of location change (context). The second type of input in ec(h)o is user movement. Each exhibit in ec(h)o is annotated with concepts that are visually represented in the exhibit (visual concepts, Section 4.4.1). For example, an exhibit with photos of pioneer explorers is annotated with a concept of 'History of Collecting'. When a user stops in a particular location (exhibit), the system interprets this as interest in the visual concept. The user model updates or adds the visual concepts as interests to the model. A set of parameters controls the influence of the visual concepts on the model.

The user model uses a spring model to keep interests balanced. The level of interest is represented by the real number and can range[10] from 0 to 10. The sum of all interests never exceeds the value of 30. In the model we consider only positive influences from the user interaction that directly increase the level of some of the interests. When this increase causes an imbalance (the sum is above 30), the implemented spring model proportionally decreases values of other interests. This mechanism supports a highly dynamic nature of the user model and guarantees that only a certain number of interests can have a high value. Another characteristic of this mechanism is that it forces the system to 'forget' the 'older' interests in favor of recently invoked interests. When the interest value drops below a set threshold during the update process the interest is removed from the model altogether.

5. Implementation

Figure 4 shows the architecture of the ec(h)o system. ec(h)o was implemented and tested in a public exhibition space at Canadian Museum of Nature in Ottawa in March 2004. The system used a combined Radio Frequency Identification (RFID) and optical sensing for position tracking. The system tracked the "x, y" coordinates of each visitor approximately every 1.6 s with a spatial resolution of 0.3 m. In terms of hardware, the position tracking system used a separate array of video cameras but all sensing data was integrated.

In addition, we used the "eyes" vision system[11] to allow for quicker refresh rates. The vision module included color video cameras connected to desktop computers to cover specified interactive zones. A camera positioned on the ceiling

---

[10] The range of the values for individual interests and their total was selected to achieve a desired proportion between object ranking criteria (see Section 4.4.2).

[11] http://www.squishedeyeballs.com

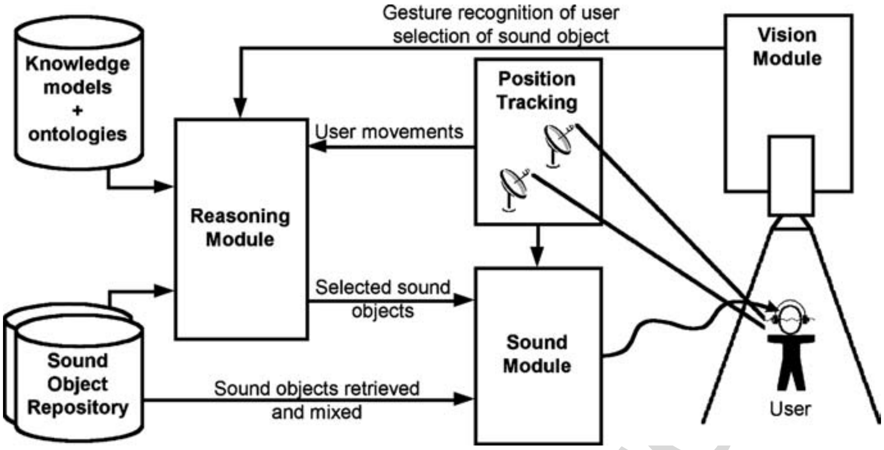

Figure 4. ec(h)o architecture.

above the artifacts was used to detect the rotation of the cube by visitors within one camera zone[12] in combination with the positioning system.

The sound module consists of a sound-file playback and mixing system driven by the position-tracking module. User position information is provided by the position tracking system and used to dynamically mix the soundscapes the user is immersed in. The sound module uses a custom-designed software mixing system implemented on a single computer. We have developed an authoring environment for mapping sounds to the physical topology of an exhibition. The delivery of the audio objects is through a stereo audio interface using FM radio transmission to portable FM receivers. In our testing environment the system served four simultaneous users. The system scales simply by adding more FM transmitters. The vision and audio delivery systems were developed in our lab using the Max/MSP environment.

The reasoning module was fully implemented with all features described in the previous sections. The real-time nature of the ec(h)o environment was the driving force for the selection of an implementation platform that supported the reasoning engine. As shown in Figure 5, the Jess inference engine is at the center of the reasoning module. We have used DAMLJessKB to load DAML+OIL ontologies into Jess (for details see Kopena and Regli, 2003). DAMLJessKB uses Jena toolkit to convert ontologies into RDF triples that are converted to Jess facts. When converted ontologies are loaded into Jess, the rules representing DAML+OIL semantics infer the missing relations in the RDF graph. This happens at start-up time and prepares the system to respond to the input in a real-time fashion. In the development version we embedded the reasoning engine in the Tomcat

---

[12] The zone for the camera depends on the height of the mount and height of the hand handling the cube. For example, the zone diameter for the camera mounted at 4 m can be as wide as 15 m with a wide angle lens.

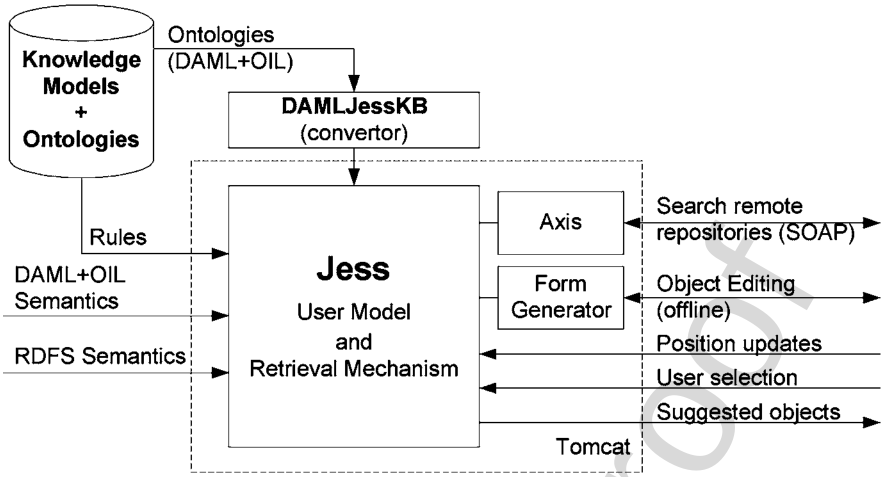

Figure 5. Implementation schema of the reasoning module.

environment in order to facilitate online editing of knowledge models as shown in Figure 5. However, for the final deployment we used the reasoning engine as a standalone application for performance reasons. All communication with the reasoning engine was accomplished through User Datagram Protocol (UDP) connections.

The user model that forms the significant part of the reasoning engine was implemented[13] using a combination of rules and specific Jess extensions via Java classes to support computation tasks such as object ranking and the spring model calculations used to compute the user interests.

We produced over 600 reusable audio objects at a low level of granularity and annotated them with the ontological information. The average length of each audio object is approximately 15 s. The shortest length is 5 s and the longest 31 s. The prefaces typically last 3 s. A majority of informational and narrative audio objects originated from the interviews with researchers and staff from the Canadian Museum of Nature in Ottawa. We subsequently scripted the objects and used actors for the recordings. For details on the content development (see Wakkary et al., 2004).

6. Evaluation

Evaluation of ubiquitous computing systems is extremely complex as these systems 'bridge the physical and online worlds' and require seamless navigation between the two, without imposing significant cognitive load on the user (Spasojevic and Kindberg, 2001). There is no agreed upon framework for evaluation of such

---

[13] The only part of the user model that was not continually updated in the final prototype was the user type as the size of the final exhibition did not provide enough supporting data for inferring this information.

systems as known in other domains such as information retrieval (trec.nist.gov) or Robocup (robocup.org). Although Burnet and Rainsford (Burnett and Rainsford, 2001) argue for a hybrid approach combining quantitative and qualitative evaluations situated in a well-defined environment, such as a ‘smart room’ (Pentland, 1996), many projects use ad-hoc evaluation approaches borrowed from other better established domains. These typically include an analysis of log files for various events and user activities, observing user behavior and conducting user interviews. The small number of test users is also an issue in that it does not allow one to make strong conclusions. For example, the evaluation of the deployment of mobile computing systems in the Exploratorium museum project provided ‘existence proofs’ for certain reactions and phenomena based on a mix of log files, observations and interviews with a small number of users rather than statistical evidence (Fleck et al., 2002).

We have found Miller and Funk’s (2001) view of the problem of evaluation of ubiquitous computing systems from the traditional ‘validation’ and ‘verification’ perspective very useful. In regard to validation, we evaluate whether the system performs the functions it was built for based on the requirements specification. Verification tests the system against the reality-checking of user evaluation to see whether the system provides the envisioned benefits.

Following Miller and Funk’s approach allowed us to focus our evaluation on areas where we researched novel approaches in adaptive ubiquitous systems. We also avoided the evaluation of aspects of the system that are not well defined or understood. Below we describe three *validation* steps for two main components of the ec(h)o system, the user model and system response:

1. User model updates: the user and environment models are updated with respect to model modifiers that represent observed user actions in the environment. The user model update mechanism interprets the meaning of the actions as conveyed by the model modifiers to adjust modeled user characteristics, i.e. in our case, the level of user interests. In the user model validation we measure how well the model changes user interests with respect to the input and interaction criteria set for ec(h)o.
2. System response: the second validation we performed evaluates how the system selects audio objects based on the user characteristics with respect to the interaction criteria.
3. User interaction: in this validation step we evaluated user interaction. We evaluated the audio objects characteristics the user selected against the interaction criteria.

In the system *verification* we obtained qualitative data that measured user experience. We developed questionnaires and performed interviews focusing on user’s perception and satisfaction with the system from the perspective of our key research questions.

6.1. VALIDATION OF THE USER MODEL FLEXIBILITY

As mentioned in Section 4.5 the rule-based user model provides a generic structure that enables the system developer to consider several inputs that influence the level of user interests in the user model. These inputs influence initial interest selection, object selection, and location change. In addition, the model allows the developer to tune the relative influence of each input using a set of parameters. The spring model implemented in the user model keeps the rest of the model balanced with the maximum values of each interest capped at a value of 10 and the sum of all interest values at 30.

Each of these actions has a different effect on the user interests. In order to achieve a well-balanced user model we designed a series of tests that evaluated how the rules responded to each type of user action. The second series of tests was designed to balance the relative influence of each type of action in the context of typical user interaction. Both tests were performed in a laboratory setting and they used variations of previously observed user interaction.

We performed a series of tests in which we tested the different combinations of parameters for the maximum interest value (maximum-concept), audio object selection contribution (primary-concept and inferred secondary-concept), location change contribution (visual-concept), and initial user interests (initial-concept). Table II shows the range of values for each parameter tested.

The goal of this test was to find a combination of parameters that would establish the dynamics in the user model with the following characteristics: moderate evolution in user interests when listening to audio objects, significant influence of changing context (visual concepts in exhibits), and protecting the user model from the domination[14] of a few concepts. Similarly, in the initiation stage we were looking for the balance between concepts initially selected by a new user and how these are combined with visual concepts when a user enters the first exhibit. It should be noted that the user model is only one component used in the ranking of audio objects; there are other factors that significantly influence object selection and overall interaction (as shown in Section 4.4.2).

Table II. Values of tested parameters

Parameter	Tested values
Initial-concept	5, 7, 10
Primary-concept	0.7, 1, 1.5, 2
Visual-concept	1, 2, 3
Maximum-concept	8, 10, 12

---

[14] As a result this would prohibit exploration of other concepts of interest and lock the user into a few concepts.

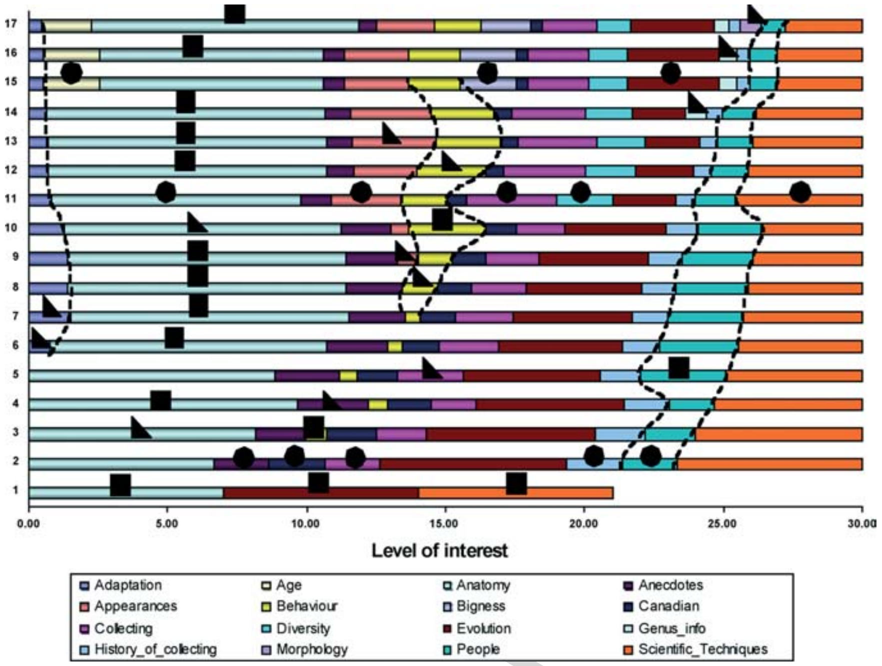

Figure 6. User model dynamics with response to user interaction with the exhibit and listening to the audio objects.

In order to simulate user interaction input we used a fixed sequence of steps that were captured from the users interacting with an earlier version of ec(ho) in our lab. We evaluated all the combinations of parameters by analyzing the graphical representations of the user model as shown in Figure 6. The figure shows the sequence of steps and evolution of interests in each step. In the first step the user selects three concepts as his or her initial concepts of interest. The circle icon indicates concepts introduced to the model by the visual concepts in the exhibit in which the user enters (Step 2, 11, and 15). In the rest of the steps the user selected audio objects. The square icon indicates a primary concept of interest and a triangle icon denotes a secondary concept of interest in the selected audio object. Figure 6 demonstrates some of the significant features of our user model. A broken line on the left shows how a concept of interest ('Adaptation') introduced to the model via listening to audio objects is being continually reduced as other concepts of interest are increasing in value. The dynamics is highlighted for the other two concepts of interest ('Behavior' in the middle and 'People' on the right). The same applies to the initial concept of interest ('Scientific techniques' furthest to the right) that is not selected and its value is reduced continuously. Figure 6 also shows how the value for a concept of interest ('Anatomy') is not increasing once it has reached the maximum value.

The user model proved to be very flexible and responsive to the parameter values and allowed us to control the dynamics of the interest levels. The combination of parameters that supported the dynamics of the user model closest to our goals was seven for the initial-concept, 1.5 for the primary-concept, two for the visual-concept, and 10 for the maximum-concept. These values supported the contribution of the user model at the level of 36% of the overall ranking of the audio objects (column 'ec(h)o' in Table I). We kept these values fixed for the rest of the evaluation.

The selected combination of parameter values is specific to our ec(h)o application and not individual users. It is likely that other applications would require different dynamics. Our model is rule-based and designed to be highly flexible. This not only allows us to modify the values of the parameters that suit the application but also to introduce new parameters into the model as needed.

6.2. USER EXPERIMENT SETUP

We installed the ec(h)o system in an existing exhibition about collecting called 'Finders and Keepers'. The exhibition contains seven exhibits, five of which are booth-type exhibits, each with several dozens of artifacts organized around topics. Two exhibits are open exhibits with larger artifacts such as a mastodon skeleton. For the exhibition, we created three interactive zones: two in booth-type exhibits and one in an open space exhibit.

The formal user evaluation included six participants. The participants had previous experience with interactive museum systems such as docent tours (three participants), interactive kiosks (3), audiotape systems (4), film and video (5), seated and ride-based systems (2) and personal digital assistant systems (2). The test group included two men and four women, aged 25–53-years old.

The testing session for each user started with a brief introduction on the purpose and testing procedure. Participants had an opportunity to interact with the system while one of the researchers accompanied them to explain how to use it. We logged all the interactions of this tutorial phase but as this was a "coached" session we did not include this data in our final evaluation. After this short training session the users had an opportunity to ask questions and seek clarification. Next, participants engaged the system as a typical museum visitor would. Users began by selecting their initial concepts of interest and they were then left alone to freely explore the exhibition. We logged all interactions with the system and used this data for the evaluation of the system described in the following sections. After the main testing session, the users were asked to complete a questionnaire. Finally, we conducted and videotaped a semi-structured interview with each participant.

In addition to the six users we tested the system with two expert reviewers. These experts included a senior researcher and senior interaction designer from

Table III. Test session characteristics

User ID	Length	#Steps	#Selections	#Locations
User1	10:36	27	19	8
User2	6:19	11	7	4
User3*	8:56	22	12	10
User4	9:53	21	16	5
User5	9:18	22	17	5
User6*	5:01	16	7	9
Expert1	15:03	32	23	9
Expert2	17:58	36	29	7

the museum. Both were familiar with the exhibit and its underlying concepts. The experts tested the system for an extended period of time with specific focus on the depth of the content and meaningfulness of the interaction. After each of the expert testing sessions we discussed the issues the experts wanted to clarify. Finally, they provided an extensive written report on the system performance.

Table III shows the characteristics of each user session: the total length of the interaction, number of interaction steps, number of selected and listened to audio objects, and number of location changes. As can be seen from Table III the number of location changes for User3 and User6 are exceptionally high. After examining the log files we found that the system repeatedly registered the single event of entering the same exhibit. This may have been caused by either the user moving along the exhibit boundary or by an error in the position-tracking module[15] . As explained in the previous section, this event caused the user model to be updated with the concepts represented by the exhibit (visual concepts), which skewed the object selection process towards those concepts. Therefore we did not include these two users in our evaluation data.

6.3. VALIDATION OF THE SYSTEM RESPONSE (OBJECT SELECTION)

In section 6.1 we showed how the system interprets the user actions and how user actions are used to update the user model and specifically the level of user interests. In this section we present our results of the recommender part of the system that selects audio objects to be offered to the user.

To evaluate the system response capabilities we have used interaction criteria. The level of fulfillment of interaction criteria can be observed from the audio objects offered to the user at each interaction step. To measure the system performance with respect to interaction criteria we defined three characteristics: variety, sustained focus, and evolution. These characteristics measure semantic relationships between offered audio objects with respect to concepts these audio objects represent.

---

[15] It is possible that participants were exploring the soundscape feature that is played when users leave an interactive zone and stops playing when users enter a zone. This starting and stopping of the soundscape would result from weaving along the boundary of an interactive zone.

In ec(h)o, at each interaction step three objects are offered O_{s1}, O_{s2}, and O_{s3}. Each object is annotated with a primary and secondary concept of interest it represents P_{s1}, P_{s2}, P_{s3} and S_{s1}, S_{s2}, S_{s3} respectively. If we define two sets $P_s = \{p|$ unique interests in $\{P_{s1}, P_{s2}, P_{s3}\}\}$, $S_s = \{s|s$ is unique interest in $\{S_{s1}, S_{s2}, S_{s3}\}\}$, and $\|M\|$ denotes the number of elements in the set M then we can define three criteria as follows:

Variety – describes the richness of choices for further interaction at each interaction step. The variability is a basic mean to put users in control of selecting topics of further interaction. It also compensates for an inherent inaccuracy of user interest modeling by providing multiple alternatives. Formally, we define variety in interaction step s as Var_s

$$Var_s = c_1^* \|P_s\| + c_2^* \|S_s - P_s\}\|$$

where we set $c_1 = 1$ and $c_2 = 0.5$. In case of ec(h)o Var_s can range from $\langle 0, 4.5 \rangle$ so we scaled it to $\langle 0, 1 \rangle$ for a clearer comparison.

Sustained focus – An ability of the system to sustain the focus on particular interests. Mono-topical systems provide a maximum degree of sustained focus but do not follow shifting user's interests. On the other side of the spectrum are systems selecting topics randomly where the sustained focus cannot be reasonably evaluated.

$$Sust_{s+1} = c_1^* \|P_{s+1} \cap P_s\| + c_2^* \|P_{s+1} \cap S_s\| + c_2^* \|S_{s+1} \cap P_s\| + c_3^* \|S_{s+1} \cap S_s\|$$

where we set $c_1 = 1$, $c_2 = 0.5$ and $c_3 = 0.25$. In case of ec(o) $Sust_{s+1}$ can range from $<0, 6.75>$ so we scaled it to $\langle 0, 1 \rangle$ for a clearer comparison.

Evolution – An ability of the system to follow shifting user interests during interaction with the system. Adaptive systems have an ability to continually shift the focus of the interaction by continuously monitoring user's interaction. We have defined evolution as the weighted number of new concepts introduced between two steps in the interaction.

$$Evolv_{s+1} = c_1^* \|P_{s+1} - (P_s \cup S_s)\| + c_2^* \|S_{s+1} - (P_s \cup S_s)\|$$

where we set $c_1 = 1$ and $c_2 = 0.5$. In case of ec(h)o $Evolv_{s+1}$ can range from $\langle 0, 4.5 \rangle$ so we scaled it to $\langle 0, 1 \rangle$ for a clearer comparison.

Table IV shows the values of the proposed evaluation characteristics when applied to the mockup data. The rows labeled as primary and secondary concepts represent concepts of interest for three hypothetical audio objects offered to the user. The values in columns 1–10 were chosen to show how different combinations of concepts affect the three measurements. As defined above, variety is measured for each interaction step and has a value of 0 if all concepts are identical (e.g. column 1) and a value of 1 if all values are unique (e.g. column 11). The sustained focus in a particular interaction step is based on the values in this step and the previous step. The sustained focus measures how many concepts from the previous

Table IV. Behaviour of sustained focus, evolution and variety

	1	2	3	4	5	6	7	8	9	10	11	12	13
Primary concept 1	a	a	b	b	b	c	c	d	f	g	i	i	i
Primary concept 2	a	a	b	b	b	c	c	d	f	g	j	j	q
Primary concept 3	a	a	b	b	b	c	c	d	f	g	k	l	r
Secondary concept 1	a	a	b	c	c	b	d	e	d	h	l	i	q
Secondary concept 2	a	a	b	c	c	b	d	e	d	h	m	m	r
Secondary concept 3	a	a	b	c	c	b	d	e	d	h	n	p	p
Sustained Focus		1.00	0.00	0.67	0.56	0.44	0.44	0.22	0.22	0.00	0.00	0.48	0.26
Evolution		0.00	1.00	0.33	0.00	0.00	0.33	0.33	0.67	1.00	1.00	0.11	0.67
Variety	0.00	0.00	0.00	0.14	0.14	0.14	0.14	0.14	0.14	0.14	1.00	0.86	0.71

step are repeated in the next step. This information is weighted differently for primary and secondary concepts being repeated as either primary or secondary concepts (columns 7, 8, and 9 demonstrate this clearly). The evolution is also computed from the current and previous interaction step and it captures how many new concepts were introduced at the primary and secondary levels. It is not a mere complement of the sustained focus as can be seen from columns 6, 7, 8, and 9. Finally, the last three columns in Table IV shows a more realistic distribution of concepts in offered audio objects. In column 12 concepts 'i' and 'j' are repeated as primary concepts, concept 'm' is repeated as a secondary concept and concept 'i' is repeated also as a secondary concept. In column 13 two concepts are repeated: 'i' as a primary and 'p' as a secondary concept.

We have calculated the sustained focus, evolution and variety for each user interaction. Figure 7 shows actual results for one user (graphs for other users show the same trends). The horizontal axis represents interaction acts that trigger object selection. These can be either the user entering the exhibition zone (step number is circled) or the user making a selection of an audio object. When a user enters the space three new audio objects are offered. After a user makes a selection the selected object is replaced with the new one and possibly a non-selected object is replaced if it had already been offered three times.

In Figure 7 we can observe that the system supports high variety of objects in each step without significant changes between the interaction steps. However, trend lines for sustained focus and evolution demonstrate significant changes at the steps representing a change of the exhibit zone. In these points the sustained focus factor decreases significantly indicating that objects offered in the new location represent new topics of interest from those offered in the previous location. This system behavior reflects our selection of the weights established in Section 6.1, specifically the weight for visual-concept, giving a strong influence of the context on the user model. Once the user stays in the same interaction zone the sustained focus increases reflecting continual changes in the user model. The trend changes in the evolution characteristics are caused by the same decision.

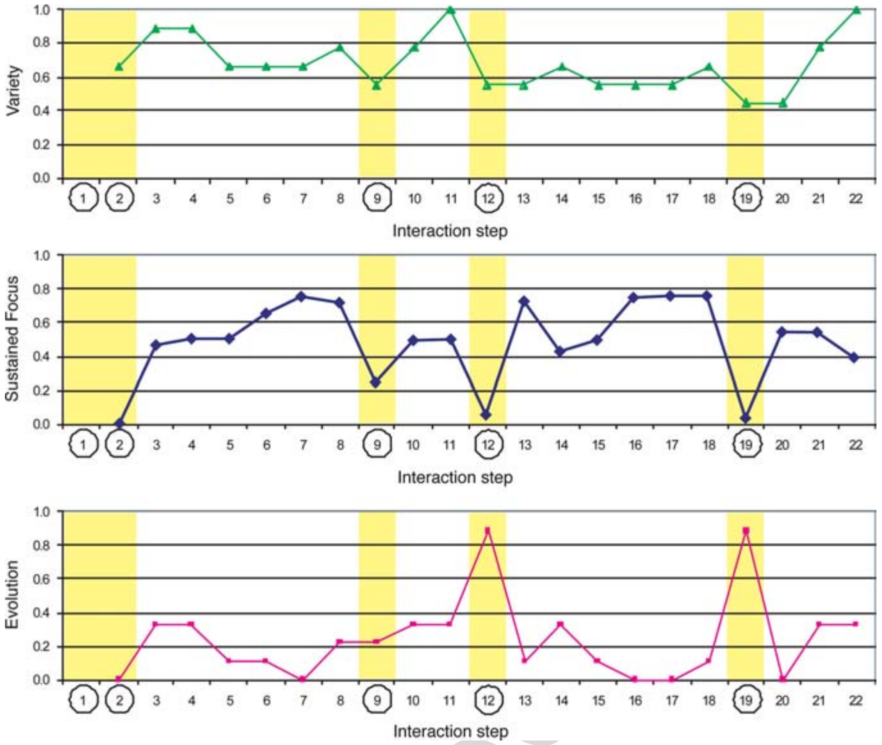

Figure 7. Variety, sustained focus and evolution for User 5.

Because the changes in the exhibit location caused such significant differences we separated the statistical processing for the ‘location-change’ steps from the ‘object-selection’ steps. Table V shows the statistical values for all three characteristics as obtained from seven test subjects.

Based on the values in Table V we can conclude that the system offers highly variable objects when users change location and the variety increases as users continue the interaction in a particular location. The high variety during the object selection steps is supported while the system maintains the focus on the concepts of interest as expressed in the user model. The low value of evolution during the object selection stage indicates the continual change in topics offered corresponding to the modest changes in the user model.

This behavior matches our expectations. As described in Section 5.3.2, several ranking criteria are combined to select audio objects offered in the next step. It is the weight with which these criteria contribute to the object ranking that determines the combination of the concepts of interest in the objects offered. To achieve different behavior from the system the relative weight of contributing criteria would have to be altered.

Table V. Statistical values for variety, sustained focus and evolution

	Overall		Selection		Location	
	AVG	STD	AVG	STD	AVG	STD
Variety	0.73	0.18	0.77	0.16	0.61	0.14
Sustained focus	0.50	0.21	0.58	0.15	0.23	0.15
Evolution	0.30	0.24	0.23	0.16	0.53	0.24

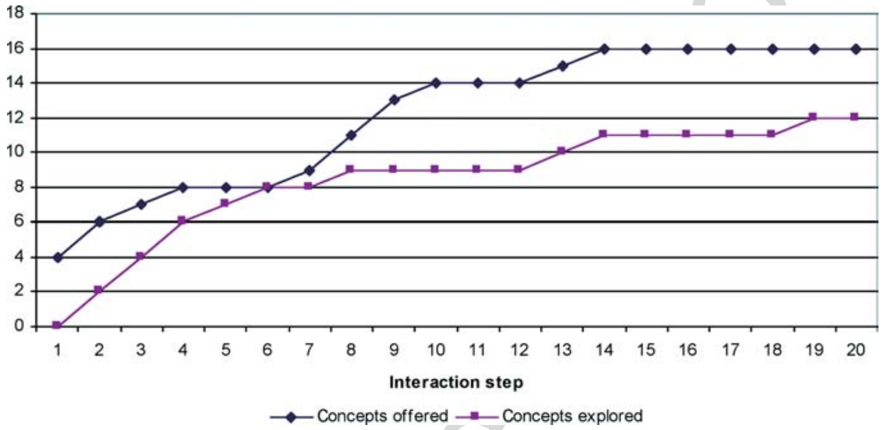

Figure 8. Concepts offered and concepts selected for User 5 (cumulative).

6.4. EVALUATION OF USER INTERACTION

While in the previous section we evaluated the system's ability to respond in the manner corresponding with our interaction criteria, in this section we examine how users interacted with the system.

As presented in the previous section we have tuned the system to favor sustained focus over evolution. However, a high level of variety enabled users to 'defy' the sustained focus of the interaction by selecting audio objects with newly introduced concepts of interest. Figure 8 shows the dynamics of how the system introduced new concepts of interest and how the users explored those concepts via object selection in the course of interaction. The horizontal axis represents interaction steps and vertical axis represents a cumulative number of concepts of interest introduced and explored up to that interaction step. The grayed areas represent steps where users changed location. The zero value for the number of concepts selected in step one is due to the fact that users did not select any object before moving into another exhibit.

The graphs in Figure 8 shows that at the beginning the system introduces new concepts at a more rapid pace. At the same time the user explores objects (and concepts of interest) rapidly until a point is reached where the user explores some

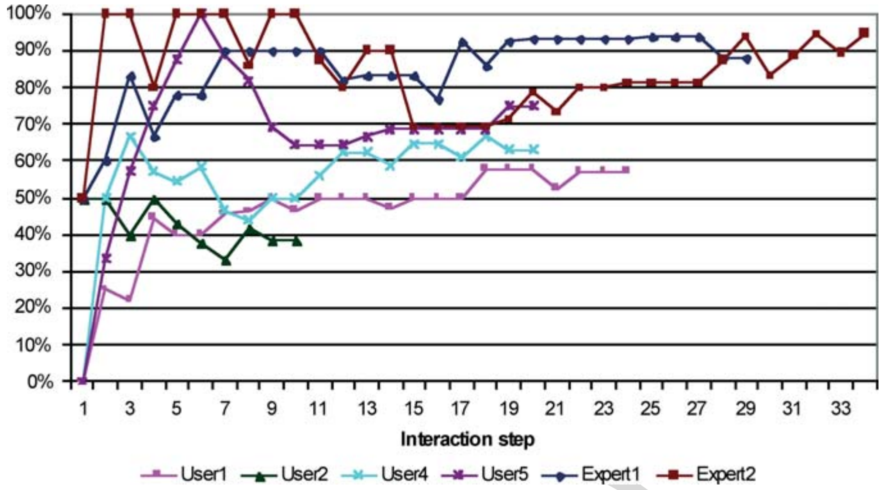

Figure 9. Concept exploration relative to concepts introduced.

of the concepts in more depth (Steps 8–12 and 14–20). Although the absolute values differ between users we have found a similar pattern is present for all users.

Figure 9 shows the percentage of selected concepts of interest by individual users relative to the number of concepts of interest introduced via offered audio objects. The graph shows that after initial steps users quickly converge to a stable proportion of the selected concepts of interest in the range of 30–70% of concepts offered.

It is difficult to speculate whether with ongoing interaction the level of concept exploration the users reached would remain at a constant level. Theoretically, as the number of concepts of interest in our system is limited to 39 concepts, the users have an opportunity to explore all of them. On the other hand, as we can see from the available data, users tend to explore certain concepts in more detail.

6.5. VERIFICATION: EVALUATION OF USER EXPERIENCE

User experience was evaluated through observation during the sessions, a questionnaire, and a semi-structured interview. The questionnaire included sixty-three questions that assessed user experience related to the overall reaction to the system, the user interface, learning how to use the system, perceptions of the system's performance, the experience of the content, and degree of navigation and control. The questionnaire also provided for open-ended written comments. Majority of the questions were on a Likert scale. Throughout the questionnaire, and especially during the semi-structured interviews we looked for an overall qualitative assessment of the experience based on Bell's ecological components of *liminality* and *engagement* (Bell, 2002). For a summary of the questionnaire results see Table VI.

Table VI. Summary of the questionnaire results on user experience (n=6; 63 questions on Likert scale of 1–5 (being best)).

Categories	Average	Standard deviation
Overall reaction (five questions including "terrible-wonderful; difficult-easy".	3.60	0.78
Tangible user interface (seven questions including "uncomfortable-comfortable; difficult-easy to manipulate; annoying-enjoyable"	4.24	0.50
Headset (two questions including "comfortable-uncomfortable to wear"	2.92	0.12
Learning curve for the system (eight questions including "difficult-easy to get started; risky-safe to explore features; unclear-clear feedback)	4.07	0.36
Perception of system performance (eight questions including "slow-fast system response; never-always reliable"	3.83	0.39
Quality of the content (fifteen questions including: "uninformative-informative; generalized-customized for me; rigid-playful; predictable-surprising"	3.78	0.52
Quality of the audio experience (nine questions including "confusing-clear; mechanical-human-like; wasteful-valuable"	3.67	0.30
Navigation and control (eight questions including "never-always able to navigate in an efficient way; always-never found myself lost in the system; always-never found myself uncertain of system state	3.23	0.29

Participants found the system enjoyable and stimulating, perhaps in part due to its novelty. The general sense of satisfaction was split between those participants who liked the playful approach and those who did not. While our sample was small we noted a clear age difference in that the "younger" participants rated satisfaction higher based on their liking of the playful approach (this was confirmed in the semi-structured interviews).

Among the factors that stood out as most positive for the participants was that the cube and audio delivery were seen as playful. The open-ended written comments and semi-structured interviews made this point clear as well. The tangible user interface was well received especially in terms of ergonomics and ease of use. This was not a surprise to us since our early testing and participatory design sessions provided us with considerable feedback, especially on ease of use and enjoyment. We went through several iterations of the wooden cube selecting the lightest wood we could find (balsa wood) and going through several form factors tested against different hand sizes. This may have also resulted in the fact that learning to use the interface and navigation were rated highly and participants felt the system had a low learning curve and so it was easy to get started. It should be stated that we provided a short tutorial on the system at the beginning of each evaluation (see Section 6.2) but nevertheless this feedback is encouraging. Interestingly, the audio content was perceived to be both accurate and clear. The issue of trust and delivery style is an area to further investigate. Since we collected the information directly from scientists and staff at the museum rather than a more generic source we wonder if this contributed in part to this result (Wakkary et al., 2004). These results lead us to believe that the system meets or satisfies many of the current advances of museum guide systems.

The questionnaire did point out challenges and areas for further research. Some things we expected such as the headphones were uncomfortable, yet to such a degree that we are currently rethinking the tradeoff between personalized spatial audio and use of headphones. Other results point to a threshold in the balance between levels of abstraction and local information. Since visitors had difficulties at time connecting what they were listening to and what was in front of them (in part this was an inherent challenge in the exhibition since the display cases had dozens to over a hundred artifacts). In many respects this confirms our finding that the ontological approach did not provide a clear enough contextual link between the artifacts and the audio information. In addition, we see both a threshold point in play versus focused attention on the exhibit in that the question relating to the content asking if it was "distractive-synergistic" scored 2.83. This raises the issue of balance in play and the possibility to shift attention away from the environment rather than play as a means of further exploring the environment.

In an open-ended question in the questionnaire and through the semi-structured interviews we explored the issues of liminal play and engagement. The results here are quite clear that play was a critical experiential factor in using the system. It was often remarked how the experience was similar to a game:

> *"At first it felt a little bit strange, especially holding this cube that looked like a children's toy, and I felt a little bit awkward about doing that, but I got over that pretty quickly. The whole system to me felt a lot like a game. I mean I got lost in it, I found myself spending a lot [more] time in a particular area then I normally would. And just the challenge of waiting to hear what was next, what the little choice of three was going to be. Yeah … So I found it over all engaging, it was fun, and it was very game-like."* (Participant 5)

The playfulness did in most instances suggest a quality of engagement that led to learning even through diverse types of museum visits including the visitor who browses through quickly but is still looking to be engaged, to the repeat visitor who experiences the audio information differently each time:

> *"I learned a lot and well you know I'm a scientist here, and I think anybody going through, even people who are in a real rush, are going to pick up some interesting facts going through. And … I mean, that was good, the text was great and was short enough that somebody in a rush is still going to catch the whole thing. And there wasn't much delay really, I mean once you showed your cube it came up pretty fast, and that is important with museum-goers. I think museum-goers don't stand and spend*

> *a bunch of time in one spot so it has to be something that comes up pretty quickly."* (Participant 4)

As mentioned earlier, there is a threshold between play in support of the exhibit on display and play with the system that can be an end in itself and even a distraction. For example, one participant occasionally focused more attention on playing with the system than the exhibition due to her enthusiasm for the game-like quality. In addition, people respond to play differently and can be argued to belong to different types of players (Bartle, 1990). One participant would have preferred a more serious and "non-playful" approach. In this case the playfulness and short length of the audio was seen as anecdotal rather than serious and scholarly.

7. Discussion

At the outset of this paper we acknowledged the challenge to capture the larger context through user modeling, particularly in ubiquitous and mobile computing applications. No doubt Fischer poses the problematic as a description of an ongoing research program than a question that a single project can address (Fischer, 2001). Nevertheless, our strategies along this front included the sensing and inference based on visitor movement, like many other systems, however, we also utilized a mixed criteria, combining ranking of concepts of interest based on direct user selection of audio objects *mixed* with visual concepts that we mapped to the context (see Sections 4.3.2 and 6.1). Our aim here was to allow for the possibility of new interests to form externally through the context. As it turned out, in analyzing the participants' selections of audio objects based on the interaction criterion of *evolution* (see Section 6.3), significant changes occurred less through user selection (this was always possible since we maintained high degree of variability in concepts at all times) than from visitors moving to another exhibit. The criterion of *evolution* can be said to evaluate internal influences (user's reflection on content) and external influences (user's reflection on context). This was possible given our aim to consider user interest as dynamic and evolving based on the interaction with the environment. In fact, we earlier stated that we do not see our system as a museum guide, recommending things based on what people like or know at the outset of a visit, rather we see it as a way to provide enrichment to the ongoing experience of the exhibit and artifacts.

The specific problem we stated at the outset of the paper was how to support the fuller experience design goals as well as functionality with an integrated modeling technique and use of semantic technologies in combination with an audio augmented reality and tangible user interface approach.

In regard to functionality, the user experience results show that ec(h)o was extremely easy to use and quick to learn, and the overall system performed well (see Section 6.5). The validation of the ec(h)o components, namely user model and object selection, showed that these performed at the required level of accuracy and flexibility. While we did not perform a comparative test with other systems, in the verification it was clear that participants had experience with many different museum based systems (see Section 6.2) and we can expect that comparisons were made with past experiences in evaluation of ease-of-use, learning curve, and performance.

In regard to the experience design goals of play and discovery, we feel our integrated modeling approach implemented two techniques to facilitate wider exploration and the discovery of new topics of interests and the ability to make new connections among topics and artifacts. The first being the aim of keeping interests balanced such that a given topic or set of topics does not dominate and prevent exploration of new topics, for this we used a spring model to proportionately moderate levels of interest (see Section 6.1). As we stated, it is important that the user model learns to "forget older interests" so that newer ones can be invoked. The second technique is to maintain a high level of variety of primary and secondary interests among the objects presented. This affords greater opportunity for the user to evolve his or her interest through a reflection on content as discussed above (see Section 6.3). These techniques contribute to the goal of establishing dynamics in the user model that support exploration and discovery of new interests through moderating evolution in the user interests, maintaining significant influence of changing context (when a visitor moves to another exhibit), and protecting against the domination of a few concepts that would choke off exploration.

We introduced the evaluation of system response or in our case, object selection based on interaction criteria of *variety*, *sustained focus*, and *evolution*. We've found these terms useful in the discussion above and we can say that we can measure *variety*, and rationalize it together with *evolution* as dependent factors in exploration and discovery of new user interests through interaction. *Sustained focus* is less clear of a measure at this stage and something we will investigate in future research.

There are cautions in our findings. The first is designers must strike a balance or they run the risk of users engrossed in the playing with the system at the expense of interacting with their surroundings, as one participant commented happened to her periodically. The second caution stems from the results that indicate that visitors had difficulties at times connecting what they were listening to and what was in front of them. It may be that the system did not always provide a coherent story, a resulting tradeoff due to its dynamic nature. Nevertheless, a much richer model of discourse and storytelling could be an option to pursue. In addition, users in the museum settings are significantly connected with concrete artifacts while ec(h)o experimented with the idea of the connection between artifacts and audio objects residing at a higher ontological level. The results indicate that either a much richer model is needed or the hypothesis of linking objects at higher abstract ontological levels is not suitable for ubiquitous context-aware applications or it has to be combined with other approaches.

8. Conclusion and Future Work

ec(h)o is an augmented audio reality system for museum visitors that was developed and tested for the Canadian Museum of Nature in Ottawa. In ec(h)o we tested the feasibility of audio display and a tangible user interface for ubiquitous computing systems – one that encourages an experience of play and engagement. The interface uses audio as the only channel to deliver short audio objects that can originate from the web. We have built several ontologies that richly described the museum environment and artifacts, audio objects and user interests. The knowledge-based recommender system builds a dynamic user model based on user choices and user movement through the exhibit and recommends audio objects to the user.

The findings of this project are positive while also calling for more research in several areas. First, we found that it is possible to build a highly flexible and accurate user model and recommender system built on information observed from user interaction that supports play and discovery as well as functionality. Ontologies and rule-based approaches proved to be a strong combination for developing such systems, yet some museum visitors are looking for more coherent stories that are highly contextualized. The ontological approach did not prove satisfactory and either more extensive knowledge engineering is needed or it has to be combined with stronger narration or discourse models. As museums are highly social places, another area that needs more research is extending the system with support for groups and group interaction.

Acknowledgements

Work presented in this paper is supported by a Canarie Inc. grant under the E-Content program. Authors would especially like to thank Dr. Mark Graham and his colleagues at the Canadian Museum of Nature in Ottawa for their enthusiastic support of this project. We would also like to thank the participants of our several workshops that contributed to the development of the project, and namely our colleagues and students, Kenneth Newby, Dale Evernden, Leila Kalantari, Doreen Leo, Gilly Mah, Robb Lovell, Mark Brady, Jordan Williams and Phil Thomson for their contributions.

References

Alfaro, I., Zancanaro, M., Cappalletti, A., Nardon, M. and Guerzoni, A.: 2003, Navigating by Knowledge. *In: Proceedings of the Eighth International Conference on Intelligent User Interfaces,* Miami, Florida, USA, pp. 221–223.

Aoki, P. M., Grinter, R. E., Hurst, A., Szymanski, M. H., Thornton, J. D. and Woodruff, A.: 2002, Sotto Voce: Exploring the Interplay of Conversation and Mobile Audio Spaces. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* Minneapolis, Minnesota, USA, pp. 431–438.

Aoki, P. M. and Woodruff, A.: 2000, Improving Electronic Guidebook Interfaces using a Task-Oriented Design Approach. *In: Proceedings of the Conference on Designing Interactive Systems,* New York City, New York, USA, pp. 319–325.

Bartle, R. A.: 1990, Who Plays MUAs?. *Comms Plus!* 18–19.

Baus, J., Krüger, A. and Wahlster W.: 2002, A Resource-Adaptive Mobile Navigation System. *In: Seventh International Conference on Intelligent User Interfaces,* San Francisco, California, USA, pp. 15–22.

Bell, G. and Kaye, J. J.: 2002, Designing technology for domestic spaces: a kitchen manifesto. *Gastronomica – the Journal of Food and Culture* **2**(2), 42–62.

Bell, G.: 2002, Making Sense of Museums: The Museum as Cultural Ecology. *Intel Labs* 1–17.

Benelli, G., Bianchi, A., Marti, P., Not, E. and Sennari, D.: 1999, HIPS: Hyper-Interaction within the Physical Space. *IEEE Multimedia Systems'99*, Florence, Italy, pp. 1075–1078.

Berners-Lee, T., Hendler, J. and Lassila, O.: 2001, The Semantic Web. *Scientific American* **284**(5), 34–44.

Bjork, S., Holopainen, J., Ljungstrand, P. and Akesson, K. P.: 2002, Designing Ubiquitous Computing Games: A Report from a Workshop Exploring Ubiquitous Computing Entertainment. *Personal and Ubiquitous Computing* **6**, 443–458.

Bobrow, D. G.: 1991, Dimensions of interactions. *AI Magazine* **12**(3), 64–80.

Bodker, S.: 1990, Through the Interface: A Human Activity Approach to User Interface Design. Malwah, New Jersey: Lawrence Erlbaum Associates, Inc.

Brewster, S., Lumsden, J., Bell, M., Hall, M. and Tasker, S.: 2003, Multimodal Eyes-Free Interaction Techniques for Wearable Devices. *In: Proceedings of the conference on Human factors in computing systems*, Ft. Lauderdale, Florida, USA, pp. 473–480.

Burke, R.: 2002, Hybrid Recommender Systems: Survey and Experiments. *User Modeling and User-Adapted Interaction* **12**(4), 331–370.

Burnett, M. and Rainsford, C.: 2001, A Hybrid Evaluation Approach for Ubiquitous Computing Environments. *Ubicomp01 Workshop: Evaluation Methodologies for Ubiquitous Computing,* http://zing.ncsl.nist.gov/ubicomp01/papers/Ubiquitous_Computing_Evaluation.pdf. Last viewed Aug. 15, 2005.

Carroll, J. M.: 2002, Human-Computer Interaction in the New Millenium. New York: ACM Press.

Carroll, J. M.: 2000, Making use: Scenario-Based Design of Human-Computer Interactions. Cambridge, Massachusetts: MIT Press.

Carse, J. P.: 1987, Finite and Infinite Games. New York: Ballantine Books.

Chandrasekaran, B., Josephson, J. R. and Benjamins V. R,: 1991, What are ontologies, and why do we need them? *IEEE Intelligent Systems* **14**(1), 20–26.

Crampton-Smith, G.: 1995, The Hand that Rocks the Cradle. *ID Magazine*, 60–65.

Crofts, N., Doerr, M. and Gill, T.: 2003, The CIDOC Conceptual Reference Model: A standard for communicating cultural contents, *Cultivate Interactive*, 9, 7 February 2003, http://www.cultivate-int.org/issue9/chios/. Last viewed Aug. 15, 2005.

Csikszentmihalyi, M.: 1990, *Flow: The Psychology of Optimal Experience*. Harper & Row, New York.

Dean, D.: 1994, *Museum Exhibition: Theory and Practice*. Routledge, London.

Dourish, P.: 2001, *Where the Action is: The Foundations of Embodied Interaction*. MIT Press, Cambridge, Massachusetts.

Dourish, P.: 2004, What We Talk About When We Talk About Context. *Personal Ubiquitous Computing* **8**(1), 19–30.

Eckel, G.: 2001, Immersive Audio-Augmented Environments. In: *Proceedings of the Eighth Biennial Symposium on Arts and Technology at Connecticut College*, New London, CT, USA. http://viswiz.gmd.de/~eckel/publications/eckel01a.pdf. Last viewed Aug. 15, 2005.

Ehn, P.: 1989, *Work-Oriented Design of Computer Artifacts*. Arbetslivscentrum, Stockholm

Fink, J. and Kobsa, A.: 2002, User Modeling for Personalized City Tours. *Artifical Intellegence Review* **18**(1), 33–74.

Fischer, G.: 2001, User Modeling in Human\– Computer Interaction'.*User Modeling and User-Adapted Interaction* **11**(1–2), 65–86.

Fishkin, K. P.: 2004, A Taxonomy for and Analysis of Tangible Interfaces. *Personal Ubiquitous Computing* **8**(5), 347–358.

Fitzmaurice, G. W., Ishii, H. and Buxton, W. A. S.: 1995, Bricks: Laying the Foundations for Graspable User Interfaces. *Proceedings of the SIGCHI conference on Human factors in computing systems*, Denver, Colorado, USA, pp. 442–449.

Fleck, M., Frid, M., Kindberg, T., O'Brien-Strain, E., Rajani, R. and Spasojevic, M.: 2002, From Informing to Remembering: Ubiquitous Systems in Interactive Museums. *IEEE Pervasive Computing* **1**(2), 13–21.

Gay, G. and Hembrooke, H.: 2004, *Activity-Centered Design: An Ecological Approach to Designing Smart Tools and Usable Systems*. MIT Press, Cambridge, Massachusetts.

Goecks, J. and Shavlik, J.: 2000, Learning Users' Interests by Unobtrusively Observing their Normal Behavior. In: *Proceedings of the fifth international conference on Intelligent user interfaces*. New Orleans, Louisiana, USA, pp. 129–132.

Goßmann, J. and Specht, M.: 2002, Location Models for Augmented Environments. *Personal and Ubiquitous Computing* **6**(5–6), 334–340.

Hatala, M., Wakkary, R. and Kalantari, L.: 2005, Rules and Ontologies in Support of Real-Time Ubiquitous Application. *Web Semantics: Science, Services and Agents on the World Wide Web* **3**(1), 5–22.

Hatala, M., Kalantari, L., Wakkary, R. and Newby, K.: 2004, Ontology and Rule Based Retrieval of Sound Objects in Augmented Audio Reality System for Museum Visitors. In: *Proceedings of the 2004 ACM symposium on Applied computing,* Nicosia, Cyprus, pp. 1045–1050.

Hinckley, K., Pausch, R., Goble, J. C. and Kassell, N. F.: 1994, Passive Real-World Interface Props for Neurosurgical Visualization. In: *Proceedings of the SIGCHI conference on Human factors in computing systems,* Boston, Massachusetts, USA, pp. 452–458.

Holmquist, L. E., Redström, J. and Ljungstrand, P.: 1999, Token-Based Access to Digital Information. *First International Symposium on Handheld and Ubiquitous Computing,* Karlsruhe, Germany, pp. 234–245.

Huizinga, J.: 1964, *Homo Ludens: A Study of the Play-element in Culture*. Beacon Press, Boston.

Hummels, C. and Helm, A.v.d.: 2004, ISH and the Search for Resonant Tangible Interaction. *Personal Ubiquitous Computing* **8**(5), 385–388.

Ishii, H., Wisneski, C., Orbanes, J., Chun, B. and Paradiso, J.: 1999, Pingpongplus: Design of an Athletic-Tangible Interface for Computer-Supported Cooperative Play. *Conference on Human Factors in Computing Systems*, Pittsburgh, Pennsylvania, USA, pp. 394–401.

Ishii, H. and Ullmer, B.: 1997, Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms. *Conference on Human Factors in Computing Systems,* Atlanta, Georgia, USA, 234–241.

Konstan, J. A., Rield, J., Borchers, A. and Herlocker, J.L.: 1998, Recommender Systems: A GroupLens Perspective. *Recommender Systems: Papers from the 1998 Workshop,* Menlo Park, California, USA, pp. 60–64.

Kopena, J. and Regli, W.C.: 2003, DAMLJessKB: A Tool for Reasoning with the Semantic Web. *IEEE Intelligent Systems* **18**(3), 74–77.

Leinhardt, G. and Crowley, K.: 1998, Museum Learning as Conversational Elaboration: A Proposal to Capture, Code and Analyze Museum Talk. Vol. Museum Learning Collaborative Technical Report MLC-01. Pittsburgh, PA: University of Pittsburgh, Learning Research and Development Center.

Lieberman, H.: 1995, Letizia: An Agent that Assists Web Browsing. *International Joint Conference on Artificial Intelligence (IJCAI-95),* Montreal, Quebec, Canada, pp. 475–480.

Miller, C. and Funk, H. B.: 2001, Verification through User Value: Or 'How to Avoid Drinking Your Own Bathwater in Ubicomp Evaluations. *Ubicomp'01 Workshop: Evaluation Methodologies for Ubiquitous Computing,* http://zing.ncsl.nist.gov /ubicomp01/papers/Ubicomp-eval3.doc. Last viewed Aug. 15, 2005.

Mladenic, D.: 1996, Personal WebWatcher: Implementation and Design. Vol. Technical Report IJS-DP-7472. Department for Intelligent Systems, J.Stefan Institute.

Mulholland, P., Collins, T. and Zdrahal, Z.: 2004, Story Fountain: Intelligent Support for Story Research and Exploration. In: *Proceedings of the Ninth international conference on Intelligent user interface,* Funchal, Madeira, Portugal, pp. 62–69.

Mynatt, E. D., Back, M. and Want, R.: 1998, Designing Audio Aura. In: *Proceedings of the SIGCHI conference on Human factors in computing systems,* Los Angeles, California, USA, pp. 566–573.

Nardi, B. A.: 1995,*Context and Consciousness: Activity Theory and Human-Computer Interaction*. MIT Press, Cambridge, Massachusetts.

Nardi, B. A. and O'Day, V. L.: 1999, *Information Ecologies: Using Technology with Heart*. MIT Press, Cambridge, Massachusetts.

Norman, M. and Thomas, P.: 1990, The Very Idea: Informing HCI Design from Conversation Analysis. In: P. Luff, N. Gilbert and D. Frohlich (eds): *Computers and Conversation.* London: Academic Press, pp. 51–65.

Not, E. and Zancanaro, M.: 2000, The MacroNode Approach: Mediating between Adaptive and Dynamic Hypermedia. In: *Proceedings of the International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems*, Trento, Italy, pp. 167–178.

Noy, N. F. and Hafner, C. D.: 1997, The State of the Art in Ontology Design: A Survey and Comparative Review. *AI Magazine* **18**(3), 53–74.

Oppermann, R. and Specht, M.: 2000, A Context-Sensitive Nomadic Exhibition Guide.*In: Proceedings of the 2nd international symposium on Handheld and Ubiquitous Computing*, Bristol, UK, pp. 127–142.

Pentland, A.: 1996, Smart Rooms. *Scientific American* **274**(4), 68–76.

Petrelli, D., Not, E., Zancanaro, M., Strapparava, C. and Stock, O.: 2001, Modeling and Adapting to Context. *Personal Ubiquitous Computing* **5**(1), 20–24.

Pirhonen, A., Brewster, S. and Holguin, C.: 2002, Gestural and Audio Metaphors as a Means of Control for Mobile Devices. In: *Proceedings of the SIGCHI conference on Human factors in computing systems,* Minneapolis, Minnesota, USA, pp. 291–298.

Post, S. and Sage, A. P.: 1990, An Overview of Automated Reasoning. *IEEE Transactions on Systems, Man and Cybernetics* **20**(1), 202–224.

Proctor, N. and Tellis, C.: 2003, The State of the Art in Museum Handhelds in 2003'. *Museum and the Web,* Pittsburgh, Pennsylvania, U.S.A., http://www.archimuse. com/mw2003/papers/proctor/proctor.html. Last viewed Aug. 15, 2005.

Resnick, P. and Varian, H. R.: 1997, Recommender Systems. *Communications of the ACM* **40**(3), 56–58.

Semper, R. and Spasojevic, M.: 2002, The Electronic Guidebook: Using Portable Devices and a Wireless Web-Based Network to Extend the Museum Experience. *Museum and Web,* Charlotte, North Carolina, USA, http://www.archimuse. com/mw2002/papers/semper/semper.html. Last viewed Aug. 15, 2005.

Seo, Y. and Zhang, B.: 2000, A Reinforcement Learning Agent for Personalized Information Filtering. In: *Proceedings of the fifth international conference on intelligent user interfaces,* New Orleans, Louisiana, USA, pp. 248–251.

Serrell, B.: 1996, The Question of Visitor Styles. In: S. Bitgood (ed.): *Visitor Studies: Theory, Research, and Practice.* Jacksonville, Alabama: Visitor Studies Association, pp. 48–53.

Shaer, O., Leland, N., Calvillo-Gamez, E. H. and Jacob, R. J. K.: 2004, The TAC paradigm: Specifying tangible user interfaces. *Personal and Ubiquitous Computing* **8**(5), 359–369.

Sparacino, F.: 2002, The Museum Wearable: Real-Time Sensor-Driven Understanding of Visitors' Interests for Personalized Visually-Augmented Museum Experiences'. *Museums and the Web,* Boston, Massachusetts, U.S.A.,http://www.archimuse.com/mw2002/papers/sparacino/sparacino.html. Last viewed Aug. 15, 2005.

Spasojevic, M. and Kindberg, T.: 2001, Evaluating the CoolTown User Experience.*Ubicomp'01 Workshop: Evaluation Methodologies for Ubiquitous Computing,* http://zing.ncsl.nist.gov/ubicomp01/papers/Exploratorium-UbicompPosition.pdf. Last viewed Aug. 15, 2005.

Suchman, L. A.: 1987, *Plans and Situated Actions: The Problem of Human-Machine Communication*. Cambridge University Press, New York, NY.

Terrenghi, L. and Zimmermann, A.: 2004, Tailored Audio Augmented Environments for Museums. In: *IUI '04: Proceedings of the ninth international conference on Intelligent user interface,* Funchal, Madeira, Portugal, pp.334–336.

Tolmie, P.: 2002, Unremarkable Computing. In: *Proceedings of the SIGCHI conference on Human factors in computing systems,* Minneapolis, Minnesota, USA, pp. 399–406.

Towle, B. and Quinn, C.: 2000, Knowledge Based Recommender Systems using Explicit User Models. *Knowledge-Based Electronic Markets, Papers from the AAAI Workshop (AAAI Technical Report WS-00–04),* Menlo Park, California, USA, pp. 74–77.

Ullmer, B.: 2002, Tangible Interfaces for Manipulating Aggregates of Digital Information. PhD Thesis, Massachusetts Institute of Technology.

Ullmer, B. and Ishii, H.: 2001, Emerging Frameworks for Tangible User Interfaces'. In: J.M. Carroll (ed.): *Human-Computer Interaction in the New Millenium.* New York: Addison-Wesley, pp. 579–601.

Ullmer, B., Ishii H. and Jacob, R.J.K.: 2005, Token+constraint Systems for Tangible Interaction with Digital Information. *ACM Transactions Computer-Human Interaction* **12**(1), 81–118.

vom Lehn, D., Heath, C. and Hindmarsh, J.: 2001, Exhibiting Interaction: Conduct and Collaboration in Museums and Galleries. *Symbolic Interaction* **24**, 189–216.

Vygotsky, L. S.: 1925/1982, *Consciousness as a Problem in the Psychology of Behaviour*. Moscow: Pedagogika.

Wahlster, W. and Kobsa, A.: 1989, User Models in Dialog Systems. In: A. Kobsa and W. Wahlster (eds): *User Models in Dialog Systems.* New York, NY: Springer-Verlag, Inc.

Wakkary, R. and Evernden, D.: 2005, Museum as Ecology: A Case Study of an Ambient Intelligent Museum Guide. In: D. Bearman and J. Trant (eds): *Museums and the Web 2005 Selected Papers,* Vancouver, Canada, pp. 151–162.

Wakkary, R., Newby, K., Hatala, M., Evernden, D. and Droumeva, M.: 2004, Interactive Audio Content: The use of Audio for a Dynamic Museum Experience through Augmented Audio Reality and Adaptive Information Retrieval. In: D. Bearman and J. Trant (eds): *Museums and the Web 2004 Selected Papers, Arlington, Virginia*, pp. 55–60.

Wakkary, R.: 2005, Framing Complexity, Design and Experience: A Reflective Analysis. *Digital Creativity* **16**(1), (in press).

Weiser, M.: 1993, Some Computer Science Issues in Ubiquitous Computing. *Communications of ACM* **36** (7),75–84.

Authors' Vitae

Marek Hatala

School of Interactive Arts and Technology, Simon Fraser University, 2400 Central City, 10153 King George Highway, Surrey, BC, Canada, V3T 2W1

Dr. Marek Hatala is Associate Professor at the School Interactive Arts and Technology at Simon Fraser University and a director of the Laboratory for Ontological Research. Dr. Hatala received his MSc in Computer Science and PhD in Cybernetics and Artificial Intelligence from the Technical University of Kosice. His primary interests lie in areas of knowledge representation, ontologies and semantic web, user modeling, intelligent information retrieval, organizational learning and eLearning. His current research looks at how semantic technologies can be applied to achieve interoperability in highly distributed and heterogeneous environments, what are the social and technical aspects of building a distributed trust infrastructures, and what role user and user group modeling can play in interactive and ubiquitous environments.

Ron Wakkary

School of Interactive Arts and Technology, Simon Fraser University, 2400 Central City, 10153 King George Highway, Surrey, BC, Canada, V3T 2W1

Prof. Ron Wakkary is Associate Professor at the School Interactive Arts and Technology at Simon Fraser University. Prof. Wakkary received his BFA. from the Nova Scotia College of Art and Design, MFA from the State University of New York at Stony Brook, and is a Ph.D. candidate in Interaction Design at the University of Plymouth, U.K. The ec(h)o project serves as a case study in his thesis on an analytical framework for interaction design. His primary research interests lie in design of interactive systems in the area of ubiquitous computing including responsive environments, personal technologies and tangible user interfaces, and the study of interaction design related methods and practice.