

# Interactive Audio Content: An Approach to Audio Content for a Dynamic Museum Experience through Augmented Audio Reality and Adaptive Information Retrieval

Ron Wakkary, Kenneth Newby, Marek Hatala, Dale Evernden and Milena Droumeva  
School of Interactive Arts and Technology  
Simon Fraser University  
Surrey, BC, Canada, V3T 2W1  
Email: rwakkary@sfu.ca

## Abstract

ec(h)o is an “audio augmented reality interface” utilizing spatialized soundscapes and a semantic web approach to information. The paper discusses our approach to conceptualizing museum content and its creation as audio objects in order to satisfy the requirements of the ec(h)o system. This includes, the conceptualizing of information relevant to an existing exhibition design (an exhibition from the Canadian Museum of Nature in Ottawa). We will discuss the process of acquiring, designing and developing information relevant to the exhibition and its mapping to the requirements of adaptive information retrieval and the interaction model. The development of the audio objects is based on an audio display model that addresses issues of psychoacoustics, composition and cognition. The paper will outline the challenges and identify the limitations of our approach.

**Keywords :** augmented reality, audio, adaptive information retrieval, museum guide

## 1. Introduction

Museums are a natural laboratory in which to examine the complex nature of constructing meaning, learning from and enjoying objects and environments through interaction. Museum visits have been described as interactive, situational, social, subjective and inter-connected with the physical environment (Leinhardt and Crowley 1998; Lehn, Heath et al. 2001). Research and commercial practice in the development of electronic museum guides have typically focused on the use of portable computing devices for interaction, data storage and audio delivery (Proctor and Tellis 2003). Despite the growing use of such systems there are known limits to this approach. These include cognitive and learning difficulties of using a new graphical interface, competition for attention between the device and its surroundings, and the ergonomic problems of weight and operation (Woodruff, Aoki et al. 2001; Proctor and Tellis 2003). Often portable computing based systems deliver content in ways familiar to computing but not familiar to museum visitors (Leinhardt and Crowley 1998; Lehn, Heath et al. 2001; Woodruff, Szymanski et al. 2001). An arguably more important limitation in current practice is the approach to digital content. Typically content for museum guides has been developed much like CD-ROM content, interactive but finite and limited structurally in terms of associations and linkages.

Our goal is to design a system that fits with the interactions and everyday competencies of the museum visitor, such that it amplifies and strengthens the visitor's ability to explore, learn from and construct the meaning of exhibitions.

The paper will discuss our approach to conceptualizing the content and its creation as audio objects in order to satisfy the requirements of ec(h)o. This includes, the conceptualizing of information relevant to an existing exhibition design (an exhibition from the Canadian Museum of Nature in Ottawa).

Firstly, we will provide an overview of the ec(h)o system and interaction model, followed by a discussion of challenges in relation to adaptive information retrieval and interactive audio. Next, we will discuss the process of acquiring, designing and developing information relevant to the exhibition and its mapping to the requirements of adaptive information retrieval and the interaction model. The development of the audio objects is based on an audio display model that addresses issues of psychoacoustics, composition and cognition. The paper will outline the challenges and identify the limitations of our approach.

## **2. Context**

### **2.1 Overview of ec(h)o**

The platform for ec(h)o is an integrated audio, vision and location tracking system installed as an augmentation of an existing exhibition installation. The platform is designed to create a museum experience that consists of a physical installation and an interactive layer of three-dimensional soundscapes that are physically mapped to museum displays and the overall exhibition installation. Each soundscape consists of zones of ambient sound and "soundmarks" generated by dynamic audio data that relates to the artifacts the visitor is experiencing. The soundscapes change based on the position of the visitor in the space, their past history with viewing the artifacts, and their individual interests in relation to the museum collection. By way of a gesture-based interaction, the visitor can interact with a single artifact or multiple artifacts in order to listen to related audio information. The audio delivery is dynamic and generated by agent-assisted searches inferred by past interactions, histories and individual interests.

The source for the audio-data is digital objects. Our original sample set of digital objects was developed using content that originated from our partner museum, the Canadian Museum of Nature. In the ec(h)o context digital objects populate a network of repositories linked across different museums. The networked nature of these repositories makes it possible for visitors in the context of one museum to access data from another. For example a visitor at the Canadian museum of Nature could access content from the local repository as well as repositories of other museums or online resources.

The ec(h)o architecture consists of four main system components: position tracking, vision system, wireless audio delivery, and reasoning. Two main types of events trigger the communication between the components: user's movement through the exhibition space, and user's explicit selection of the sound objects. A more detailed description and analysis of the technical and information retrieval aspects can be found in our previous writing (Hatala, Kalantari et al.).

### **2.2 Interaction Model**

#### **2.2.1 Conversation structure**

Similar to the work of Woodruff, Aoki, Hurst and Szymanski (Woodruff, Aoki et al. 2001; Woodruff, Szymanski et al. 2001; Aoki, Grinter et al. 2002), we have adopted the storytelling structure based on Sacks' conversation analysis theory (Sacks 1974). In our case, we modeled the system and interaction

on this conversation structure. ec(h)o offers the visitor three short audio pieces that we refer to as audible icons. These audible icons serve as *prefaces*. They are in effect offering three turn-taking possibilities to the visitor. The visitor selects one and the system delivers the related audio object. This turn-taking represents the *telling* phase. After the delivery of the object, the system again offers three audible icons. It is at this stage that the *response* phase occurs. The visitor's *response* is expressed through the gesture selection with the interaction object. Additionally, the system may be met by no *response*, because the visitor does not wish to engage the system. It will then enter into a silent mode. The visitor may also have moved away and the system will initiate a soundscape and prepare for the next conversational encounter.

### **2.2.2 Navigational model**

It is important to explain the navigational model for both its novelty and simplicity, and of course for its support of the interaction. The audio objects are semantically tagged to a range of topics. At the beginning of each interaction cycle or "conversation", three topics are inferred to be more relevant than others to the visitor based on their user model, location and interaction history. Audio objects are cued representing each of the three chosen topics. Audible icons or *prefaces* related to the objects are presented to the visitor (each audible icon is differently spatialized in the audio display for differentiation), a visitor chooses one of the *prefaces* and listens to an object representative of the topic chosen. The topics are not explicit to the visitor, rather the consistency and content logic is kept in the background. After listening to the object, the visitor is offered a new *preface* based on their previous topic selection. The two previous *prefaces* that were not selected are offered once again. If after three offerings of the same *preface* and topic have transpired without selection, that topic is replaced. A more detailed description and analysis of the interaction model and design process can be found in a previous writing (Wakkary, Hatala et al.).

Figure 1. 1-2-4 navigation model

## **3. Challenges**

In related works (Bederson 1995; Sarini and Strapparava 1998; Andolesk and Freedman. 2001), the relationship of the digital content to the artifacts is either pre-planned and fixed, or the digital content is not networked and limited to the local device, in some cases both limits are true. ec(h)o employs a semantic web approach to the museum's digital content thus it is networked, dynamic and user-driven. The interface of ec(h)o does not rely on portable computing devices, rather it utilizes a combination of gesture and object manipulation recognized by a vision system.

The dynamic and user-driven nature of ec(h)o requires a highly responsive retrieval mechanism with a criteria defined by psychoacoustics, content and composition domains. The retrieval mechanism is based on a user model that is continually updated as the user moves through the exhibition and listens to the audio objects. The criteria are represented by rules operating on the ontological descriptions of sound objects, museum artifacts and user interests.

The capturing of the user interests is in the center of the research of several disciplines such as information retrieval, information filtering and user modeling (Wahlster and Kobsa 1989). Most of the systems were developed for retrieval of documents where document content is analyzed and explicit user feedback is solicited to learn or infer the user interests. In the context of ec(h)o there is no direct feedback from the user. ec(h)o can be categorized as a personalized system as it observes user's behavior and makes generalizations and predictions about the individual user based on their interactions (Goecks and Shavlik 2000; Kobsa and Fink 2002).

Particular challenges in relation to the use of audio in ec(h)o include: the designing and preparing of the audio objects for dynamic and personal delivery; the information management aspects of developing classifications and relationships; and the ultimate need to create an audio display and user experience that is coherent, consistent and pleurably exploratory in relation to an existing exhibition. The following section focuses on how we addressed these challenges.

## **4. Audio Content Design and Development**

We will describe the design and production of the audio content in four stages: 1) Our expert-based system approach to data collection, describing how we acquired the raw information related to the exhibition and artifacts; 2) Concept mapping and audio object design, discussing the initial knowledge management design of the information and the design and development of objects; 3) Design of audio objects, description of the audio display and acoustical experience issues related to the objects as audio; 4) User scenarios and inference rules, a discussion of our development of user scenarios as a design approach to developing the inference rules. This set of descriptions outlines the entire process of the design and development of the audio objects, the ontologies, and inference rules.

### **4.1 Stage One - Expert Based System Approach for Data Collection**

In order to acquire the relevant information for the audio objects we devised two modes of interviews with researchers at the Nature Museum of Canada. The interview sessions took place in the museum over the course of several days. Our goal was to develop an information gathering process that paralleled our conversation approach to the interaction model. The aim was that the interview processes would provide us with audio material that could be used directly and would create the experience for the museum visitor of an interactive guide to the museum with a group of different 'experts' (in the end we used the interview texts to create a script and so we did not use the recordings directly). In keeping with our conversation model we hoped to emulate the experience of experts conversing (both with themselves and the visitor), each taking turns contributing bits of information based on their particular interests and area of expertise.

We organized interviews with members of the museum research staff. These individuals were chosen based on their expertise in a number of different knowledge domains related to the exhibition -- Zoology, Ichthyology, Botany, Vascular Plants, Invertebrates, conservation, etc. The interviews were conducted in two parts: part one, introduced the interviewee to the ec(h)o project and asked them to comment or provide contextual information from their perspective and area of expertise related to the exhibit; phase two, involved a video walk-through of the exhibit space in which the interviewer and expert engaged in a discussion of the artifacts and collections on display. Here interviewees were asked to provide discipline specific information about the exhibits themes and sub-themes, as well as relevant information about specific artifacts within each of the exhibits. Here is a sample set of questions:

- Each display tells a story, what is that story?
- Can you discuss the different groupings of the artifacts and explain how and why they are clustered?
- Can you describe the significance of each artifact or group of artifacts?
- What makes these particular artifacts best suited to their task?
- Can you describe type of sounds that you think would supplement this exhibit?
- How might these sound effects work to enhance the visitor experience?

- Can you speak to the potential of linking content to other museums?

The results of the interviews were largely successful, however there were problems and gaps in our information set. Some interviewee's limited their discussion to high-level explanations of the exhibit that were difficult to integrate into a museum visit, while others provided interesting anecdotal information about artifacts. While we wanted to avoid an "encyclopedia" approach to the information, we supplemented the interview information gathered with research of the museums archives and research collections. We met with archival experts to filter through potential source material that already existed. Source material was, for the most part, limited to audio tracks taken from studies conducted in the field, as well as video productions that the museum had collected or produced over the years.

## **4.2 Stage Two - Concept Mapping & Ontologies**

### **4.2.1 Concept map development**

In order to translate the information gathered in the interview process for adaptive retrieval, we needed to conceptualize the information within a loose taxonomy or concept map that could eventually be developed into semantic web ontologies. The concept map would guide us in the design and relationships of the information in the form of digital objects. As part of the information management related research, the strength of a semantic web approach is the interoperability of generic and specific topic ontologies. We wanted to test the ability to develop specific ontologies that could function with generic ontologies. In addition, we were very aware that the existing curatorship and exhibition design represented a knowledge map in its own right relevant to the objects and collections on display, nevertheless our goal was to insert another level of knowledge mapping that could be productively "superimposed" over the existing exhibition.

In order to develop the concept map for the ec(h)o version of the exhibition we analysed the recorded video and audio from the expert interviews. This analysis entailed watching and listening to video and audio followed by a mapping process. This was undertaken by the entire interdisciplinary research team which helped ensure that the design of the concept map could function in the different contexts of adaptive retrieval, audio display and user experience. The concepts and themes that the team clustered were organized into a relational map. These concepts and themes became classifiers that were used during the meta-tagging stage of audio object development. Conceptual and thematic classifiers evolved out of the concept mapping exercise, where as the topical classifiers were taken from the established Dewey Decimal Classification system. The conceptual map served as an important visualization tool that helped the team understand the topical and conceptual links between artifacts, and exhibit sections. The map also served as a point of departure for helping the team recognize potential openings for bringing in content from other museums. The concept map was also the starting point for the development and adoption of different ontologies.

Figure 2. Preliminary concept mapping

### **4.2.3 Ontologies**

The interaction model is based on the semantic description of the content of the objects. We have developed an ontology where a sound object is described using several properties. As an ability to link to other museums is an important feature of ec(h)o our ontology builds significantly on the standard Conceptual Reference Model (CRM) for heritage content developed by CIDOC (Crofts, Dionissiadou et al. 2001). The CRM provides definitions and a formal structure for describing the implicit and explicit concepts and relationships used in cultural heritage documentation. To describe sound objects

we use CRM `TemporalEntity` concept for modeling periods and events and `Place` for modeling locations. We describe museum artifacts using the full CRM model.

The content of the sound object is not described directly but annotated with three entities: concepts, topics, and themes. The concepts describe the domains that are expressed by the sound object such as evolution, behaviour, lifestyle, diversity, habitat, etc. Since the collections in individual museums are different so are the concept maps describing these collections. A topic is a more abstract entity that is represented by several concepts, such as botany, invertebrates, marine biology, etc. To facilitate the mappings between topic ontologies in individual museums we have mapped the topics to the Dewey Decimal Classification whenever possible. Finally, themes are defined as entities supported by one or more topics, e.g. the theme of “bigness: in invertebrates and marine biology”.

### **4.3 Stage Three - Audio Object Design**

In this stage, the aim was to design and develop the audio objects that supported the interaction and audio display model, and that could be classified and meta-tagged based on the concept mapping. In support of the interaction model, the audio objects would need to be different types of audio elements—prefaces, audio objects, sound-marks, and keynote sounds. In the early stages of this work we focused on developing audio objects, and their corresponding prefaces—sound mark, and keynote sound production would come later. The production of the audio objects started with dividing up the interviews into manageable information objects. In doing so it was clear that each object needed to be cognitively manageable for the user, as well as manageable for the system, which meant it needed to be classifiable. Embedded references to artifacts were either made explicit or removed all together, and the scripts for all objects were edited to be suitable, as well as interesting, to as broad an audience as possible.

Once refined, each discrete audio object was then entered into the repository database where they were meta-tagged for retrieval purposes. Meta-tagged information includes location and associative information such as the exhibit an object belonged to, as well as the specific artifact it was most relevant to. Objects were also meta-tagged based on their topical, conceptual, and thematic qualities. For example, if we were classifying an object that spoke about the collecting tools used by early plant collector Catherine Parr Trail, its topical classifier would be “botany”, its conceptual classifier “tools and techniques,” and its thematic classifier, “early collectors”.

#### **4.3.2 Audio Display**

In order to deliver a seamless integrated audio display experience, ec(h)o works on several levels –the first mode of interaction involves a movement-related immersion in a dynamic soundscape, related thematically to different parts of the exhibit; another, second mode of interaction, the visitors engage with the audio display installation via a manipulation object, responding to spatially displayed audio prefaces; and a third level is knowledge acquisition or learning, by listening to the audio knowledge objects. It is important that all levels work together, physically, cognitively and psychoacoustically in order to deliver a worthwhile immersive experience. Issues of sound amplitude and frequency range must be considered for all elements of the audio display system.

In addition, we felt it was important to provide the visitor a variety of voices with a spectrum spanning the serious and authoritative to the playful and whimsical. Before recording the audio objects, consideration was given in choosing the voice talents to perform the scripted content. Issues of gender, voice quality, timbre, clarity and other psychoacoustic sound markers came into play. For ec(h)o, an even gender split between the voices is used with care in order to develop differences in both timbre and performance and to facilitate easy discrimination due to variations in range, frequency and timbre. The voices consist of one deeper, broad-range strong male voice, one warmer timbre, softer male

voice, one mid-high pure female voice, and one deeper, richer timbre female voice. For an initial database of just over 200 short sound objects, four different voices (two male and two female) appeared to be sufficient to provide the diverse, yet consistent and recognizable audio web of information.

In order to create an atmosphere of ‘engaged and fun learning,’ the aural design attempts to stay away from a highly accented ‘authoritative’ presentation of the museum information. For this reason, voice talents used are not professionals, but ‘real’ people. The style of narration determined during recording is natural pace, moderate inflection, with an even dynamic speech envelope in distinction to the emphasized polished performance typical of professional voice talents.

Preliminary testing of different approaches in the presentation of informational audio options—options that are effective in pointing to thematically or conceptually different information objects—suggests a conversational approach is appropriate to maintaining a level of playful engagement and dialogue with ec(h)o.. Since this approach is based on a style of presenting artifact information that has a ‘teasing,’ humorous quality, the vocal approach taken is appropriately different. Of the four voice talents delivering the audio objects, two are used to present the prefaces – one male and one female voice. Again, the objective is to have a natural, spontaneous voice, but with a greater emphasis on ‘character’ – a more upbeat accent and inflection. This enhances the immediate playful engagement of the museum visitor and, as a consequence of this engagement, successfully provokes a greater interest in selecting a particular audio object.

#### **4.3.3 Audio Production**

Once the scripted objects are transferred to audio, the files are compressed in high-resolution mp3 format for the purpose of quick retrieval over an Internet connection. Given the slight loss of fidelity due to this compression technique it is essential that the source recordings be clear and of optimal amplitude to from the outset, in order to be clearly heard through the transmitted wireless audio format.

Another important production process to be considered is the storing and categorizing of the database of audio objects as a basis for a cross-institutional adaptive information retrieval and interaction model system. One option for a naming convention involves using a semantic signifier in combination with a numerical index: [botany]\_00001.mp3, where this signifier could be derived from any of the subject tags applied to each individual record. The final ec(h)o audio object database design omits this signifier due to possible future inconsistencies with the collections of other museums that might wish to participate in the development and sharing of knowledge object repositories.

### **4.4 Stage Four - Inference Rules**

In order to develop the inference rules we developed three models to conceptualize and test the rules: 1) visitor model; 2) narrative model; 3) soundscape model. In addition to the content and content mapping process outlined above, we relied on our initial observational and site studies of museums and museum visitors, discussions with museum administrators, exhibition designers and curators, and the research literature in museum studies (Lehn, Heath et al. 2001; Sparacino 2002).

#### **4.4.1 visitor models**

Our visitor model is comprised of three classifications of users:

- A *busy* visitor does not want to spend much time in each exhibit; instead he/she wants to stroll through the museum to get a general idea;
- An *avaricious* visitor wants to know as much as possible; He does not rush and moves from one exhibit to another in near sequent order;

- A *selective* visitor mostly chooses sound objects that represent certain concepts.

There are three levels of interest: -1 (indicates disinterest), 0 (indicates some interest), and 1 (indicates more interest), but they can be extended. Visitor's interest are computed as follows:

- When an *avaricious* visitor enters an exhibit, and is slow, his/her interests will be asserted to the primary concept of any narration that describes an artifact in that exhibit. This makes sense, because we do not need to be picky about interests and we can assume that he is interested in almost any concept.
- Interests of a selective visitor does not get easily overwritten. The rules engine should infer new interests only after he repeatedly choose narrations with certain concepts.
- For each exhibit, we need to calculate what is primary concept of most narrations that are about that exhibit. The interests of a busy visitor can only be overwritten with those, when he enters an exhibit.
- For any visitor, when he repeatedly refuses to listen to narrations with certain primary concepts, we can infer his/her disinterest to those concepts.

#### 4.4.2 Narrative models

In addition, to our goal of linking repositories and ontologies across different museums, we also faced the task of linking content across different exhibit sections. In order to maintain coherency in an ec(h)o visitors experience we saw it as being necessary to provide meaningful links between audio objects. To facilitate this, it would be important to avoid situations where a clear disconnect existed between two audio objects. In defining the notion of a clear connection we identified the following categories of linkage types:

- *Artifact to artifact*: This occurs when the content of two audio objects makes reference to, or explicitly speaks to, the same artifact. For example: audio object A and audio object B fall into this category when they both reference the same moose antlers.
- *Concept to concept*: This link occurs when two audio objects are conceptually linked—for example, audio object A and audio object B might both talk about adaptation, and could therefore be linked with out being considered discontinuous. Note: it is our assumption that concept to-concept links are less tangible then artifact-to-artifact links. It is also worth noting that an audio object will often speak to more then one concept. When multiple concepts are present in an audio object, it is usually possible to discern one that is more prominent then the others present. Therefore, a classification hierarchy of sorts can exist when we consider the notion of an audio objects conceptual make up—that is, we might have an audio object with a primary concept, and a secondary concept. Here secondary concepts are defined as being less explicit then primary.
- *Localized links*: The notion of the localized link comes from the observation that visitors like to explore when they are taking in an exhibit. The idea here is that disconnects are not always a bad thing, and that visitors find inherent satisfaction in the experience of re-orienting themselves. To provide for this, we have made room for supporting discontinuous links between objects, as long as they are at least partially contextually localized—that is, in the same exhibit space.

Based on the above explanation, linkage classifiers were formalized and used to create rules that the system could then manage. In total two types of linkage classifiers were developed (primary, and secondary) and each classifier was given a point value. Point values reflected the concept of linkage



tangibility. It was our assumption that, in general, conceptually linked objects are less tangible (unless, of course, the concept is made very explicit) than artifact linked objects.

The primary link classifiers were those described in the discussion above (artifact, concept, and localized) Secondary link classifiers deal with the presence of contextual information embedded in an audio object itself. Context information is defined as that which makes explicit reference to an artifact—ie: “the shell marked number 5”...or, “the moose antlers in the center of the display”.

Contextual information helps to facilitate the visitor’s reference, and is thus important when dealing with artifact changes, and objects that are linked based on the localization classifier. Two kinds of secondary classifiers exist—contextualized, and non-contextualized.

In evaluating the linkage potential between two objects, sameness and difference across the primary link classifiers is considered. Contextualized content within the objects is also considered. To be linked, the sum of the primary and secondary scores must achieve a certain value. An artifact-to-artifact link is the most tangible, and therefore it is always classified as being linkable, regardless of its conceptual, and contextual information score. Note that contextualized objects that are not localized are prone to creating strong disconnects, therefore any objects that fall into this category is never allowed to be linked.

#### **4.4.3 Soundscape model**

The soundscape model is composed of zones of ambient sounds that are modulated when compared to a user’s interactions and interests. In addition, proximity to soundmarks effects the overall soundscape. The sounds are generally abstract in nature.

## **5. Evaluation**

Given the complex nature of the system and user evaluation we tested our design and development of the audio objects as we went along. User tests were performed to evaluate the interaction model, the use and style of prefaces and audio objects, the inference rules and narrative models, and a series of technical and integration tests that allowed for limited user testing of the overall system. The final prototype will be installed at the Nature Museum of Canada in March 2004, and we will then perform extensive user testing. The series of progressive testing allowed us to modify our current design and inform subsequent designs.

To date, users have found the interaction experience to be coherent and that design of prefaces and audio objects has been effective. Participants reported no significant issues around poor flow, or clunky content presentation. A consensus emerged in support of the style and flavor of the audio object prefaces, which were viewed as being entertaining and effective based on their ability to pique curiosity and motivate further interaction. For the most part, topical links between objects were better observed than conceptual links. Two characteristic behavior patterns emerged to indicate that our original concern over avoiding disconnections across linked objects may have been unwarranted. First, participants tended to jump across topically, and in doing so often encountered disparate content in their turn taking. Second, participants admitted that their impetus for choices was more in keeping with a need to satisfy their curiosity (curiosity created by the prefaces, that is). This partly countered our assumption that participants would be exercising choice based on a need to hear more information about a specific topic or concept. Both of these insights indicated that users were more inclined to approach the experience from a position of play, rather than structured, focused exploration. A welcomed result!

## 6. Future Work and limitations

Current limitation of our process is the timeliness by which audio objects are designed, meta-tagged and then tested. This mitigates open development of audio objects available for use within the network by other producers. In a related issue, the current system has a very limited implementation of the networked potential of the system.

In the areas of audio display and interaction, we will need further testing to evaluate if our minimal intervention in terms of contextual guidance is successful or not. We may find that visitors will need more explicit instructions either through audio or text. In addition, we have some concern about issues of selection and integration of the various modes of audio display and their combination as determined by an inference system. It should be ensured that the frequency range, amplitude and ambient elements from one sound layer are not interfering with the bandwidth and clarity of the other sound layer, for optimal auditory satisfaction.

Future work will lead us to researching further the complex roles of the design of audio objects, inference rules, audio display and the interaction model in creating an engaging and playfully exploratory interaction.

## References

- Andolesk, D. and M. Freedman. (2001). Artifact As Inspiration: Using Existing Collections And Management Systems To Inform And Create New Narrative Structures. Museum and the Web 2001.
- Aoki, P. M., R. E. Grinter, et al. (2002). Sotto Voce: Exploring the Interplay of Conversation and Mobile Audio Spaces. Proc. CHI 2002.
- Bederson, B. (1995). Audio Augmented Reality: a proptotype automated tour guide. Conference Companion CHI '95.
- Crofts, N., I. Dionissiadou, et al. (2001). Definition of the CIDOC object-oriented Conceptual Reference Model (version 3.2.1).
- Goecks, J. and J. Shavlik (2000). Learning user's interests by unobtrusively observing their normal behaviour. 5th International Conference on Intelligent User Interfaces, New Orleans.
- Hatala, M., L. Kalantari, et al. (2004). Ontology and Rule based Retrieval of Sound Objects n Augmented Audio Reality System for Museum Visitors. ACM Symposium on Applied Computing Conference, Nicosia, Cyprus.
- Kobsa, A. and J. Fink (2002). "User Modeling for Personalized City Tours." Artificial Intelligence Review **18**: pp. 33-74.
- Lehn, D. v., C. Heath, et al. (2001). "Exhibiting Interaction: Conduct and Collaboration in Museums and Galleries." Symbolic Interaction **24**(2): 189-216.
- Leinhardt, G. and K. Crowley (1998). Museum Learning as Conversational Elaboration: A Proposal to Capture, Code and Analyze Talk in Museums. Pittsburgh, Learning Research & Development Center, University of Pittsburgh: 24.
- Proctor, N. and C. Tellis (2003). The State of the Art in Museum Handhelds In 2003. Museum and the Web 2003.
- Sacks, H. (1974). An Analysis of the Course of a Joke's Telling in Conversation. Explorations in the Ethnography of Speaking. R. Bauman and J. Sherzer. Cambridge, Cambridge University Press: 337-353.

- Sarini, M. and C. Strapparava (1998). Building a User Model for a Museum Exploration and Information Providing Adaptive System. Proc. HYPERTEXT'98.
- Sparacino, F. (2002). The Museum Wearable: real-time sensor-driven understanding of visitors' interests for personalized visually-augmented museum experiences. Proc Museums and the Web.
- Wahlster, W. and A. Kobsa (1989). User Models in Dialog Systems. User Models in Dialog Systems. A. Kobsa and W. Wahlster. Heidelberg & Berlin, Springer Verlag.
- Wakkary, R., M. Hatala, et al. (2003). Echoing the conversational space of museums through audio augmented reality and adaptive information retrieval. *Submitted DIS 2004*.
- Woodruff, A., P. M. Aoki, et al. (2001). Electronic Guidebooks and Visitor Attention. International Cultural Heritage Informatics Meeting.
- Woodruff, A., M. H. Szymanski, et al. (2001). The Conversational Role of Electronic Guidebooks. Proc Ubicomp 2001, Atlanta, GA, Springer.

## **Acknowledgements**

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