Dynamic Path Prediction and Recommendation in a Museum Environment

Karl Grieser[†‡], Timothy Baldwin[†] and Steven Bird[†]

† CSSE
University of Melbourne
VIC 3010, Australia

‡ DIS
University of Melbourne
VIC 3010, Australia

{kgrieser,tim,sb}@csse.unimelb.edu.au

Abstract

This research is concerned with making recommendations to museum visitors based on their history within the physical environment, and textual information associated with each item in their history. We investigate a method of providing such recommendations to users through a combination of language modelling techniques, geospatial modelling of the physical space, and observation of sequences of locations visited by other users in the past. This study compares and analyses different methods of path prediction including an adapted naive Bayes method, document similarity, visitor feedback and measures of lexical similarity.

1 Introduction

Visitors to an information rich environment such as a museum, are invariably there for a reason, be it entertainment or education. The visitor has paid their admission fee, and we can assume they intend to get the most out of their visit. As with other information rich environments and systems, first-time visitors to the museum are at a disadvantage as they are not familiar with every aspect of the collection. Conversely, the museum is severely restricted in the amount of information it can convey to the visitor in the physical space.

The use of a dynamic, intuitive interface can overcome some of these issues (Filippini, 2003; Benford et al., 2001). Such an interface would conventionally take the form of a tour guide, audio tour, or a curator stationed at points throughout the museum. This research is built around the assumption that the museum visitor has access to a digital device such as a PDA and that it is possible for automatic systems to interact with the user via this device. In this way we aim to be able to deliver relevant content to the museum visitor based on observation of their movements within the physical museum space, as well as make recommendations of what exhibits they might like to visit next and why. At present, we are focusing exclusively on the task of recommendation.

Recommendations can be used to convey predictions about what theme or topic a given visitor is interested in. They can also help to communicate unexpected connections between exhibits (Hitzeman et al., 1997), or explicitly introduce variety into the visit. For the purposes of this research, we focus on this first task of providing recommendations consistent with the visitor's observed behaviour to that point. We investigate different factors which we hypothesise impact on the determination of what exhibits a given visitor will visit, namely: the physical proximity of exhibits, the conceptual similarity of exhibits, and the relative sequence in which other visitors have visited exhibits.

Recommendation systems in physical environments are notoriously hard to evaluate, as the recommendation system is only one of many stimuli which go to determine the actual behaviour of the visitor. In order to evaluate the relative impact of different factors in determining actual visitor behaviour, we separate the stimuli present into a range of predictive methods. In this paper we target the task of user prediction, that is prediction of what exhibit a visitor will visit next based on

Proceedings of the Workshop on Language Technology for Cultural Heritage Data (LaTeCH 2007), pages 49–56,
Prague, 28 June 2007.

their previous history. Language based models are intended to simulate a potentially unobservable source of information: the visitor's thought process. In order to identify the reason for the visitor's interest in the multiple part exhibits we parallel this problem with the task of word sense disambiguation (WSD). Determining the visitor's reason for visiting an exhibit allows a predictive system to more accurately model the visitor's future path.

This study aims to arrive at accurate methods of predicting how a user will act in an information-rich museum. The space focused on in this research is the Australia Gallery Collection of the Melbourne Museum, at Carlton Gardens in Melbourne, Australia. The predictions take the form of which exhibits a visitor will visit given a history of previously visited exhibits. This study analyses and compares the effectiveness of supervised and unsupervised learning methods in the museum domain, drawing on a range of linguistic and geospatial features. A core contribution of this study is its focus on the relative import of heterogeneous information sources a user makes use of in selecting the next exhibit to visit.

2 Problem Description

In order to recommend exhibits to visitors while they are going through a museum, the recommendations need to be accurate/pertinent to the goals that the visitor has in mind. Without accurate recomme ndations, recommendations given to a visitor are essentially useless, and might as well not have been recommended at all.

Building a recommender system based on contextual information (Resnick and Varian, 1997) is the ultimate goal of this research. However the environment in this circumstance is physical, and the actions of visitors are expected to vary within such a space, as opposed to the usual online or digital domain of recommender systems. Studies such as HIPS (Benelli et al., 1999) and the Equator project[1] have analysed the importance and difficulty of integrating the virtual environment into the physical, as well as identifying how non-physical navigation systems can relate to similar physical systems. For the purpose of this study, it is sufficient to acknowledge the effect of the physical environment by scaling all recommendations against their distances from one another.

The common information that museum exhibits contain is key in determining how each individual relates to each other exhibit in the collection. At the most basic level, the exhibits are simply isolated elements that share no relationship with one another, their only similarity being that they occur together in visitor paths. This interpretation disregards any meaning or content that each exhibit contains. But museum exhibits are created with the goal of providing information, and to disregard the content of an exhibit is to disregard its purpose.

An exhibit in a museum may be many kinds of things, and hence most exhibits will differ in presentation and content. The target audience of a museum is one indicator of the type of content that can be expected within each exhibit. An art gallery is comprised of mainly paintings and sculptures: single component exhibits with brief descriptions. A children's museum will contain a high proportion of interactive exhibits, and much audio and visual content. In these two cases the reason for visiting the exhibit differs greatly.

Given the diversity of information contained within each exhibit and the greater diversity of a museum collection, it can be difficult to see why visitors only examine certain exhibits during their tours. It is very difficult to perceive what a visitor's intention is without constant feedback, making the problem of providing relevant recommendations a question of predicting what a visitor is interested in based on characteristics of exhibits the visitor has already seen. The use of both physical attributes and exhibit information content are used in conjunction in an effort to account for multiple possible reasons for visiting as exhibit. Connections between physical attributes of an exhibit are easier to identify than connections based on information content. This is due to the large quantity of information associated with each exhibit, and the difficulty in determining what the visitor liked (or disliked) about the exhibit.

In order to make prediction based on a visitor's history, the importance of the exhibits in the visitors path must be known. This is difficult to obtain directly without the aid of real-time feedback from

[1] http://www.equator.ac.uk

the user themselves. In an effort to emulate the difficulty of observing mental processes adopted by each visitor, language based predictive models are employed.

3 Resources

The domain in which all experimentation takes place is the Australia Gallery of the Melbourne Museum. This exhibition provides a history of the city of Melbourne Melbourne, from its settlement up to the present day, and includes such exhibits as the taxidermised coat of Phar Lap (Australia's most famous race horse) and CSIRAC (Australia's first, and the world's fourth, computer). The Gallery contains enough variation so that not all exhibits can be classified into a single category, but is sufficiently specialised to offer much interaction and commonality between the exhibits.

The exhibits within the Australia Gallery take a wide variety of forms, from single items with a description plaque, to multiple component displays with interactivity and audio-visual enhancement; note, for our purposes in experimentation, we do not differentiate between exhibit types or modalities. The movement of visitors within an exhibition can be restricted if the positioning of the exhibits require visitors to take a set path (Peponis et al., 2004), which can alter how a visitor chooses between exhibits to view. In the case of the Australia Gallery, however, the collection is spread out over a sizeable area, and has an open plan design such that visitor movement is not restricted or funnelled through certain areas and there is no predetermined sequence or selection of exhibits that a given visitor can be expected to spend time at.

We used several techniques to represent the different aspects of each exhibit. We categorised each exhibit by way of its **physical attributes** (e.g. size) and taxonomic information about the **exhibit content** (e.g. clothing or animal). We also described each exhibit by way of its **physical location** within the Australia Gallery, relative to a floorplan of the Gallery.

The Melbourne Museum also has a sizable web-site[2] which contains much detailed information about the exhibits within the Australia Gallery. This data is extremely useful in that it provides a rich vocabulary of information based on the content of each exhibit. Each exhibit identified within the Australia Gallery has a corresponding web-page describing it. The information content of an exhibit is made up of the text in its corresponding web-page combined with its attributes. By having a large source of natural language information associated with the exhibit, linguistic based predictive methods can more accurately identify the associations made by visitors.

The dataset that forms that basis of this research is a database of 60 visitor paths through the Australia Gallery, which was collected by Melbourne Museum staff over a period of four months towards the end of 2001. The Australia Gallery contains a total of fifty-three exhibits. This data is used to evaluate both physical and conceptual predictive methods. If predictive methods are able to accurately describe how a visitor travels in a museum, then the predictive method creates an accurate model of visitor behaviour.

Exhibit components can be combined to form a description for each exhibit. For this purpose, the Natural Language Toolkit [3] (Bird, 2005) was used to analyse and compare the lexical content associated with each exhibit, so that relationships between exhibits can be identified.

4 Methodology

Analysis of user history as a method of prediction (or recommendation) has been examined in Chalmers et al. (1998). Also discussed is the role that user history plays in anticipating user goals. This approach can be adapted to a physical environment by simply substituting in locations visited in place of web pages visited. Data gathered from the paths of previous visitors also forms a valid means of predicting other visitors' paths (Zukerman and Albrecht, 2001). This approach operates under the assumption that all visitors behave in a similar fashion when visiting a museum. However visitors' goals in visiting a museum can differ widely. For example, the goals of a student researching a project will differ to those of a family with young children on a weekend outing.

[2] `http://www.museum.vic.gov.au/`

[3] `http://nltk.sourceforge.net/`

A conceptual model of the exhibition space is created by visitors with a specific task in mind. Interpretation of this conceptual model is key to creating accurate recommendations. The building of such a conceptual model takes place from the moment a visitor enters an exhibition, until the time they leave, and skews the visitor towards groups of conceptual locations and categories.

The representation of these intrinsically dynamic models is directly related to the task the visitor has in mind. Students will form a conceptual model based around their course requirements, children around the most visually attractive exhibits, and so forth. This necessitates the need for multiple exhibit similarity measures, however in the absence of express knowledge of the 'type' of each visitor in the sample data, a broad-coverage recommendation system that functions best in all circumstances is the desired goal. It is hoped that in future, reevaluation of the data to classify visitors into broad categories (e.g. information seeking, entertainment seeking) will allow for the development of specialised models tailored to visitor types.

The models of exhibit representation we examine in this research are exhibit proximity, text-based exhibit information content, and exhibit popularity (based on the previous visitor data provided by the Melbourne Museum), as well as combinations of the three. Exhibit information content is a two part representation: primarily each exhibit has a large body of text describing the exhibit drawn from the Melbourne Museum website. It is fortunate that this information is curated, and managed from a central source, so that inconsistencies between exhibit information are extremely rare. The authors were unable to find any contradictory information in the web-pages used for experimentation, as may be the case with larger non-curated document bodies. The second component of the information content is a small set of key terms describing the attributes of the exhibit. Textual content as a means of determining exhibit similarity has been analysed previously (Green et al., 1999), both in terms of keyword attributes and bodies of explanatory text.

In order to form a prediction about which exhibit a visitor will next visit, the probability of the transition of the visitor from their current location to every other exhibit in the collection must be known. Prediction of the next exhibit by proximity simply means choosing the closest not-yet-visited exhibit to the visitor's current location. In terms of information content, each exhibit is related to all other exhibits to a certain degree. To express this we use the attribute keywords as a query to find the exhibit most similar. We use the attribute keywords associated with each document to search the document space of the exhibits to find the exhibit that is most similar to the exhibit the visitor is currently located at. To do this we use a simple tf·idf scheme, using the attribute keywords as the queries, and the exhibit associated web pages as the document space. The score of each query over each document is normalised into a transitional probability array such that $\sum_j P(q|d_j) = 1$ for a query (q) over the j exhibit documents (d_j).

In order to determine the popularity of an exhibit, the visitor paths provided by the Melbourne Museum were used to form another matrix of transitional probabilities based on the likelihood that a visitor will travel to an exhibit from the exhibit they are currently at. I.e. for each exhibit e an array of transitional probabilities is formed such that $\sum_j P(e|c_j) = 1$ where $c_j \in C' = C/\{e\}$, i.e. all exhibits other than e. In both cases Laplacian smoothing was used to remove zero probabilities.

The methods of exhibit popularity and physical proximity are superficial in scope and do not extend into the conceptual space adopted by the visitors. They do however give insight into how a physical space affects a visitors' mental representation of the conceptual areas associated with specific exhibit collections, and are more easily observable. Visitor reaction to exhibit information content is harder to observe and more problematic to predict. Any accurate recommender systems produced in this fashion will need to take into account the limitations these two methods place on the thought processes of visitors.

Connections that visitors make between exhibits are more fluid, and are harder to represent in terms of similarity measures. Specifically it is difficult to see why visitors make connections between exhibits as there can be multiple similarities between two exhibits. To this end we have equated this problem with the task of Word Sense Disambiguation (WSD). The path that a visitor takes can be seen as a sentence of exhibits, and each exhibit in the sentence has an associated meaning. WSD is used to determine the meaning of the next exhibit based on the meanings of previous exhibits in the path. For each word in the keyword set of each exhibit, the WordNet (Fellbaum, 1998) similarity is calculated against each word in another exhibit. The similarity is the sum of the WordNet similarities between all attribute keywords in the two exhibits (K_1, K_2), normalised over the length of both keyword sets:

$$\frac{\sum_{k_1 \in K_1} \sum_{k_2 \in K_2} WNsim(k_1, k_2)}{|K_1||K_2|}$$

For the purposes of this experiment we have chosen to use three WordNet similarity/relatedness measures to simulate the conceptual connections that visitors make between exhibits. The Lin (Lin, 1998) and Leacock-Chodorow (Leacock et al., 1998) similarity measures and the Banerjee-Pedersen (Patwardhan and Pedersen, 2003) relatedness measures were used. The similarities were normalised and transformed into probability matrices such that $\sum_j P_{WNsim}(e|c_j) = 1$ for each next exhibit c_i. The use of WordNet measures is intended to simulate the mental connections that visitors make between exhibit content, given that each visit can interpret content in a number of different ways.

The history of the visitor at any given time is essential in keeping the visitor's conceptual model of the exhibit space current. The recency of a given exhibit within a visitor's history is inversely proportional to how long ago the exhibit was encountered.

To take into account the visitor history, the collaborative data, proximity, document vectors, and conceptual WordNet similarity, we adapt the naive Bayes approach. The conditional probabilities of each method are combined along with the temporal recency of an exhibit to produce a predictive exhibit recommender. The resultant recommendation to a visitor can be described as follows:

$$\hat{c} = \arg\max_{c_i} P(c_i) \sum_{j=1}^{t} P(A_j|c_i) \times 2^{-(t-j+1)} + \frac{2^{-t}}{t}$$

where t is the length of the visitor's history, $A_j \in C$ is an exhibit at time j in the visitor history (and C is the full set of exhibits), and $c_i \in C' = C/\{A_j\}$ is each unvisited exhibit. The most probable next exhibit ($\hat{c}$) is selected from all possible next exhibits (c_i). Any selections made must be compared against the visitor's history. In this, we assume that a previously visited exhibit has already been seen, and hence should not be recommended again.

The effectiveness of these methods was tested in multiple combinations, both with history modeling and without (only the exhibit the visitor is currently at is considered). Testing was carried out using the sixty visitor paths supplied by the Melbourne Museum. For each method two tests were carried out:

- Predict the next exhibit in the visitor's path.
- Only make a prediction if the probability of the prediction is above a given threshold.

Each path was analysed independently of the others, and the resulting recommendations evaluated as a whole. The measures of precision and recall in the evaluation of recommender systems has been applied effectively in previous studies (Raskutti et al., 1997; Basu et al., 1998). In the second test precision is the measure we are primarily concerned with: it is not the aim of this recommender system to predict all elements of a visitor's path in the correct order. The correctness of the exhibits predicted is more important than the quantity of the predictions the visitor visits, hence only exhibits predicted with a (relatively) high probability are included in the final list of predicted exhibits for that visitor.

The thresholds are designed to increase the correctness of the predictions, by only making a prediction if there is a high probability of the visitor travelling to the exhibit. As all predictive methods choose the most probable transition from all possible transitions, the transition with the highest probability is always selected. The threshold values simply cut off all probabilities below a certain value.

5 Results and Evaluation

The first tests carried out were done only using the simple probability matrices described in Section 4, and hence only use the information associated with the visitor's current location and not the entirety of their history. The baseline method being used in all testing is the naive method of moving to the closest not-yet-visited exhibit.

Method	BOE	Accuracy
Proximity (baseline)	0.270	0.192
Popularity	0.406	0.313
Tf·Idf	0.130	0.018
Lin	0.129	0.039
Leacock-Chodorow	0.116	0.024
Banerjee-Pedersen	0.181	0.072
Popularity - Tf·Idf	0.196	0.093
Popularity - Lin	0.225	0.114
Popularity - Leacock-Chodorow	0.242	0.130
Popularity - Banerjee-Pedersen	0.163	0.064
Proximity - Tf·Idf	0.205	0.084
Proximity - Lin	0.180	0.114
Proximity - Leacock-Chodorow	0.220	0.151
Proximity - Banerjee-Pedersen	0.205	0.105
Proximity - Popularity	0.232	0.129

Table 1: Single exhibit history using individual and combined transitional probabilities

In order to prevent specialisation of the methods over the training data (the aforementioned 60 visitor paths), 60 fold cross-validation was used. With the path being used as the test case removed from the training data at each iteration.

The results of prediction using only the current exhibit as information can be seen in Table 1. Combinations of predictive methods are also included to add physical environment factors to conceptual similarity methods. For example, if two exhibits may be highly related conceptually but on opposite sides of the exhibit space, a visitor may forgo the distant exhibit in favour of a closer exhibit that is slightly less relevant.

Due to the lengths of the recommendation sets made for each visitor (a recommendation is made for each exhibit visited), precision and recall are identical. The measure of Bag Of Exhibits (BOE) describes the percentage of exhibits that were visited by the visitor, but not necessarily in the same order as they were recommended. The BOE measure is the same as measuring precision and recall for the purposes of this evaluation. With the introduction of thresholds to improve precision, precision and recall are measured as separate entities.

As seen in Table 1 the performance of the conceptual or information similarity methods (the tf·idf method, Lin, Leacock-Chodorow and Banerjee-Pedersen) is worse than that of the methods based on static features of the exhibits, and all perform worse than the baseline. In order to produce a higher percentage of correct recommendations, thresholds were introduced. Using thresholds, a recommendation is only made if the probability of a visitor visiting an exhibit next is above a given percentage. The thresholds used in Table 2 are arbitrary, and were arrived at after experimentation.

It is worth noting that in both tests, with and without thresholds, the method of exhibit popularity based on visitor paths is the most successful. One expects this trend to continue with the introduction of the history based model described in Section 4. Each transitional probability matrix was used in conjunction with the history model, the results of this experimentation can be seen in Table 3.

Only single transitional probability matrices are used in conjunction with the history model. The physical distance to an exhibit is only relevant to the current prediction, the distance travelled in the past from exhibit to exhibit is irrelevant, and so physical conceptual combinations are not necessary. A model such as this describes the evolution of a thought process, or is able to identify the common conceptual thread linking the exhibits in a visitor's path. This is only true *if* the visitor has a conceptual model in mind when touring the museum. Without the aid of a common information thread, conceptual predictive methods based on exhibit information content will always perform poorly.

6 Discussion

The visitor paths supplied by the Melbourne Museum represent sequential lists of exhibits, and each visitor is a black box travelling from exhibit to exhibit. It is this token vs. type problem that does not allow us to select an appropriate predictive method with which to make recommendations. Instead a broad coverage method is necessary. Use of history models to analyse entire visitor paths are less successful than analysis of solely the current location of the visitor. This can be attributed to the fact that a majority of the visitors tracked may not have had preconceived tasks in mind when they entered the museum space, and just moved from one visually impressive exhibit to the next. The visitors do not consider their entire history as being relevant, and only take into account their current

Method	Threshold	Precision	Recall	F-score
Proximity	0.03	0.271	0.270	0.270
Popularity	0.06	0.521	0.090	0.153
Tf·Idf	0.06	0.133	0.122	0.128
Lin	0.01	0.129	0.129	0.129
Leacock-Chodorow	0.01	0.117	0.117	0.117
Banerjee-Pedersen	0.01	0.182	0.180	0.181
Popularity - Tf·Idf	0.001	0.176	0.154	0.164
Popularity - Lin	0.0005	0.383	0.316	0.348
Popularity - Leacock-Chodorow	0.0005	0.430	0.349	0.385
Popularity - Banerjee-Pedersen	0.001	0.236	0.151	0.184
Proximity - Tf·Idf	0.001	0.189	0.174	0.181
Proximity - Lin	0.0005	0.239	0.237	0.238
Proximity - Leacock-Chodorow	0.0005	0.252	0.250	0.251
Proximity - Banerjee-Pedersen	0.0005	0.182	0.180	0.181
Proximity - Popularity	0.001	0.262	0.144	0.186

Table 2: Single exhibit history predictive methods using thresholds

Method	BOE	Accuracy
Proximity	0.066	0.0
Popularity	0.016	0.0
Tf·Idf	0.033	0.0
Lin	0.064	0.0
Leacock-Chodorow	0.036	0.0
Banerjee-Pedersen	0.036	0.0

Table 3: Entire visitor history predictive methods.

context. This also explains the relative success of the predictive method built from analysis of the visitor paths, presenting a marked improvement over the baseline of nearest exhibit. In the best case (as seen in Table 2) the exhibit popularity predictive method was able to give relevant recommendations 52% of the time.

The interaction between predictive methods here is highly simplified. The assumption made is that all aspects of the visitor's conceptual model are independent, or only interact on a superficial level (see the lower halves of Tables 1–2). More complex methods of prediction need to be explored fully take into account the interaction between predictive methods.

Representations based on physical proximity take into account little of how a visitor conceptualises a museum space. They do however describe the fact that closer exhibits are more visible to visitors, and are hence more likely to be visited. Proximity can be used as an augmentation to a conceptual model designed to be used within a physical space.

Any exhibit is best described by the information it contains. Visitors with a specific task in mind when entering an exhibition already have a pre-initialised conceptual model, relating to a theme. The visitors seek out content related to their conceptual model, and separate the bulk of the collection content from the information they require. The representation of the content within each exhibit as a vocabulary of terms allows us to find similarity between exhibits. The data available at the time of this testing does not make the distinction between user types, and so only broad coverage methods result in a improvements.

With the introduction of user types to the data supplied by the museum, specific predictive methods can be applied to each individual user. This additional information can be significantly beneficial as the specialisation of predictive types to visitors is expected to produce much more accurate predictions and recommendations. Currently the only method available to discern the user type is to analyse the length of time the visitor spends at each each exhibit. This data is yet to be adapted and annotated from the raw data supplied by the Melbourne Museum.

7 Conclusion

The above methods are intended to represent baseline components of possible conceptual models that represent how a visitor is able to selectively assess the dynamic context of museum visits. The model that a visitor generates for themselves is unique, and is difficult to represent in terms of physical attributes of exhibits.

Being able to predict future actions of a user within a given environment allows a recommender system to influence a user's choices. Key to the prediction of future actions, is the idea that a user has a conceptual model of how they see content within the environment in relation to a task. With respect to a museum environment, the majority of users have no preconceived conceptual model upon entering an exhibition and must build one as they explore the environment. Users with a preconceived task will more often than not stick to exhibits surrounding a particular theme. Use of a language-based conceptual model based on the information contained within an exhibit can be combined with conceptual models based on geospatial attributes of the exhibit to create a representation of how a user will react to an exhibit. The use of heterogeneous information contained within the exhibit space is only relevant when the visitor has an information-centric task in mind.

7.1 Future Work

The methods dealing with a language-based conceptual model given here are very basic, and the overall accuracy and precision of the recommender system components require improvement. Additional annotation of the paths of visitors to the museum will enable proper evaluation of conceptual information based predictive methods. On-site testing of predictive methods at the Melbourne Museum is the ultimate goal of this project, and testing the effects of visitor feedback on recommendations will also be analysed. In order to gain more insight into visitor behaviour, the current small-scale set of visitors needs to be expanded to include multiple visitor types, as well as tasks.

Acknowledgments

This research was supported by Australian Research Council DP grant no. DP0770931. The authors wish to thank the staff of the Melbourne Museum for their help in this study. Special thanks goes to Carolyn Meehan and Alexa Reynolds for their gathering of data, and helpful suggestions throughout this study. Thanks also goes to Ingrid Zukerman and Liz Sonenberg for their input on this research.

References

Chumki Basu, Haym Hirsh, and William Cohen. 1998. Recommendations as classification: Using social and content-based information in recommendation. In *Proceedings of the National Conference of Artificial Intelligence*, pages 714–720, Madison, United States.

Giuliano Benelli, Alberto Bianchi, Patrizia Marti, David Sennati, and Elena Not. 1999. HIPS: Hyper-Interaction within Physical Space. In *ICMCS '99: Proceedings of the IEEE International Conference on Multimedia Computing and Systems*, volume 2, page 1075. IEEE Computer Society.

Steve Benford, John Bowers, Paul Chandler, Luigina Ciolfi, Martin Flintham, Mike Fraser, Chris Greenhalgh, Tony Hall, Sten-Olof Hellstrom, Shahram Izadi, Tom Rodden, Holger Schnadelbach, and Ian Taylor. 2001. Unearthing virtual history: using diverse interfaces to reveal hidden worlds. In *Proc Ubicomp*, pages 1–6. ACM.

Steven Bird. 2005. NLTK-Lite: Efficient scripting for natural language processing. In *Proceedings of the 4th International Conference on Natural Language Processing (ICON)*, pages 11–18, Kanpur, India.

Matthew Chalmers, Kerry Rodden, and Dominique Brodbeck. 1998. The Order of Things: Activity-Centred Information Access. *Computer Networks and ISDN Systems*, 30:1–7.

Christiane Fellbaum, editor. 1998. *WordNet: An Electronic Lexical Database*. MIT Press, Cambridge, USA.

Silvia Filippini. 2003. Personalisation through IT in museums: Does it really work? Presentation at ICHIM 2003.

Stephen J. Green, Maria Milosavljevic, Robert Dale, and Cecile Paris. 1999. When virtual documents meet the real world. In *Proc. of WWW8 Workshop: Virtual Documents, Hypertext Functionality and the Web*.

Janet Hitzeman, Chris Mellish, and Jon Oberlander. 1997. Dynamic generation of museum web pages: The intelligent labelling explorer. *Archives and Museum Informatics*, 11(2):117–115.

Claudia Leacock, Martin Chodorow, and George A Miller. 1998. Using corpus statistics and WordNet relations for sense identification. *Computational Linguistics*, 24(1):147–65.

Dekang Lin. 1998. Automatic retrieval and clustering of similar words. In *(CoLING)-(ACL)*, pages 768–774, Montreal, Canada.

Siddharth Patwardhan and Ted Pedersen. 2003. Extended gloss overlaps as a measure of semantic relatedness. In *International Joint Conference on Artificial Intelligence*, pages 805–810, Acapulco, Mexico.

John Peponis, Ruth Conroy Dalton, Jean Wineman, and Nick Dalton. 2004. Measuring the effect of layout on visitors' spatial behaviors in open plan exhibition settings. *Environment and Planning B: Planning and Design*, 31:453–473.

Bhavani Raskutti, Anthony Beitz, and Belinda Ward. 1997. A feature-based approach to recommending selections based on past preferences. *User Modelling and User Adaption*, 7(3):179–218.

Paul Resnick and Hal R Varian. 1997. Recommender systems. *Commun. ACM*, 40(3):56–58.

Ingrid Zukerman and David W Albrecht. 2001. Predictive statistical models for user modeling. *User Modeling and User-Adapted Interaction*, 11(1–2):5–18.