Story Fountain: Intelligent support for story research and exploration

Paul Mulholland, Trevor Collins & Zdenek Zdrahal
Knowledge Media Institute
The Open University, UK
+44 1908 {654506, 655731, 654512}
{p.mulholland, t.d.collins, z.zdrahal}@open.ac.uk

ABSTRACT

Increasingly heritage institutions are making digital artifacts available to the general public and research groups to promote the active exploration of heritage and encourage visits to heritage sites. Stories, such as folklore and first person accounts form a useful and engaging heritage resource for this purpose. Story Fountain provides intelligent support for the exploration of digital stories. The suite of functions provided in Story Fountain together support the investigation of questions and topics that require the accumulation, association or induction of information across the story archive. Story Fountain provides specific support toward this end such as for comparing and contrasting story concepts, the presentation of story paths between concepts, and mapping stories and events according to properties such as who met whom and who lived where.

Keywords

Web-based interfaces, Personalization and customization of interfaces, Ubiquitous interfaces and smart environments, Ontologies, Intelligent exploration.

INTRODUCTION

A prevailing aim of many heritage institutions is to allow the public to actively engage with the subject matter, answering their own questions and making their own interpretations. This draws on the long-standing mission of many heritage institutions to provide a source of education [16] and reflects contemporary approaches to learning such as constructivism that stress the importance of the active learner [9]. Many heritage institutions are investigating how new technology can help them to realize these goals.

The work described in this paper was conducted as part of the European Union funded CIPHER project, a 30 month research project that began in 2002 with the aim of supporting communities of interest in the active exploration of heritage resources, by means of story sharing [13]. Within the CIPHER project we are developing and testing easy-to-use tools for the exploration and description of digital heritage and the construction, presentation and sharing of stories in communities of interest associated with a heritage institution. Here we wish to focus on the support provided in the project for the exploration of a corpus of digital stories.

In designing support for exploration we made two assumptions. First, we characterize exploration as question driven rather than unguided. By this we mean that the person exploring the resources is generally not randomly finding and using the resources but is trying to answer one or more questions and formulate further questions for investigation. Second, we assume that the questions being explored are open-ended and evaluative rather than satisfiable solely by factual information. To support this kind of exploration we need to provide functionality for different ways of organizing and traversing the story archive that can be matched to the task and characteristics of the user. This is beyond what can directly be supported by a search engine, which primarily supports information location within pages rather than reasoning across pages.

Our work has been motivated by needs identified while working with Bletchley Park, the wartime code breaking center and home of Colossus, the world's first programmable computer. The code breaking center closed in 1945 and did not become a heritage institution until the early 1990's after decades of secrecy. Since then a group of volunteers associated to the Park have been piecing together much of its hidden history. One important source of information has been interviews conducted with those that worked there and their families. Approximately 100 interviews have been collected. Other sources of information include releases from the UK Public Records Office, a government office that holds documents until their publication is no longer considered sensitive. The interviews, Public Records Office releases and other sources each provide part of the jigsaw as to what happened in the code breaking center.

Much of our work has been in participation with the Bletchley Park Guides community. These are approximately 40 volunteers who provide tours of the site, conduct research, give lectures, inform curatorial decisions at the Park, and publish books and other educational resources. We have already developed with

them a community portal for sharing these resources in the Guides' community, plus day-to-day announcements relevant to their work (see figure 1) [6].

Figure 1. The Bletchley Park Guides portal.

Our current aim has been to provide intelligent support for the activities of the Guides community and also provide dynamically generated, customized interfaces to the tools for other audiences such as members of the general public who visit the Park. Story Fountain, the tool developed for this purpose, uses ontologies of the domain and the heritage resources plus a reasoning engine for the selection and organization of resources. Input and presentation modules provide interfaces appropriate for different audiences.

In this paper we present a review of the related work, a detailed description of the design of Story Fountain and illustrate how Story Fountain is used in two user scenarios. We also describe the system architecture and explain how the system has been implemented. After discussing the work done and the near-term future work, the paper concludes with a summary.

RELATED WORK

A number of tools have been developed that could support the pursuit of high-level, open-ended, evaluative questions using a collection of digital resources. When reviewing the literature, we have in mind Nicholas Paley's proposal that a work of art should be appreciated as a "rhizomatic assembly of densely entangled crossroads, passages, galleries and heterogeneities, complemented by the method of multiple analytic stances" [15, page 12]. Basically, by this he means that an artefact is associated with many different concepts which allow different pathways to be made within and across artworks. The viewer will take different pathways depending on the

question they are trying to answer and the question-answering methods they are applying.

From a story context, this proposal strongly relates to aspects of narrative theory. Boje [2] describes Story Network Analysis in which a social network is constructed across a set of stories, according to the characters that the stories have in common. This combines a corpus of stories into a single representation bringing into close proximity events in which a particular character participated regardless of which story mentioned this event. This allows the reader to traverse the corpus by following different characters of their choosing, rather than the structure provided by individual stories. Story Network Analysis therefore also allows for multiple pathways across content but is specifically tied to characters and stories rather than the more general proposal of Paley. We will now consider some existing technologies in the light of the above theories, drawing out lessons for the design of Story Fountain.

Work in the area of *visualization* provides a number of suggestions as to what is required for the rhizome-like exploration of heritage resources. Schneiderman [18] describes StarField displays for the visualization of a large number of items within a single representation. The display can be modified dynamically by the viewer, modifications (for example using sliders to broaden or narrow the range of data being displayed) provide immediate feedback to the user. The user-specification of the dimensions according to which the data will be displayed allows a large amount of data to be viewed within a chosen context. Hochheiser and Schneiderman [10] demonstrated how such technology could be used for exploring the works of a musician. Important lessons that have informed the design of Story Fountain are first the advantage of allowing a user to organize and view a story-base from a chosen perspective. Second, clear advantages can be seen for allowing the user to interact directly with a representation of the story-base in order to explore new perspectives.

Walden Paths [19] is another tool that can be used to support exploration across a set of digital resources. Walden Paths is aimed at school level education and allows a teacher to specify an exploratory path for a student through a set of digital resources. The student can follow the teacher-specified path and at any point leave the path to explore on their own. The student can then easily return to the path when they wish. This work illustrates the potential importance of path generation and following when investigating a story corpus.

Wrappers, such as those described by Muslea [14] are used to extract information from pre-defined templates and compose them into a new digital document that summarizes from a large number of sources. Wrappers are used for example to construct news information, hotel availability and transport timetables. Wrappers, and

technologies that provide contextualized search results [7] indicate the usefulness of extracting selected information from sources. From our perspective, embedding story summaries within a representation of a set of stories can provide an effective overview and facilitate its use.

Brusilovsky [3] introduces a number of *personalization technologies* that tailor the interface according to the needs of different audiences. Three kinds of personalization are identified: content, presentation and navigation. Content personalization is when the subject matter is selected or filtered according to, for example, the educational attainment reached by the user in previous sessions with the application. Similarly, presentation and navigation can be tailored according to audience requirements. Within a tool for story exploration the interface should also ideally be modified according to different content, presentation and navigation policies depending on the goals, skills and experience of the user.

Finally, within a heritage context some tools have been developed to support resource exploration. One interesting example is *HyperMuseum* [20]. This tool allows a user to take different paths through digital artifacts. HyperMuseum supports exploration through the use of WordNet [8] to find similar concepts that can be searched for. In Story Fountain we wish to support exploration by means of an explicit conceptual structure of the stories and domain represented in an ontology.

In summary, based on a review of the literature the design of Story Fountain was guided by the following requirements:

- Allow organization of a set of stories according to a chosen perspective.

- Provide an interactive representation of a set of stories to allow continued, uninterrupted exploration.

- Support the finding and following of paths through the story-base.

- Embed story summaries within representations of a story set to provide an effective overview and facilitate its use.

- Provide different input and presentation modules for content, navigation and presentation appropriate to the audience.

- Support exploration using an explicit conceptual structure of stories and the domain.

DESIGN

Story Fountain was designed to support the pursuit of open-ended, high-level questions using a corpus of digital stories. These are questions where the answer is between the stories rather than in a particular story, and answering requires judgment and interpretation by the reader.

Ontologies

Story Fountain uses ontologies to describe the stories and the domain to which they relate. The basic conceptual structure of the narrative and story ontology is shown in figure 2. In the ontology, a distinction is made between the story (what is told) and the narrative (how it is told). This draws on the work of Chatham [4]. A narrative presents a story, and also has a presentation style and URI. A single story may have multiple narratives, for example, for different audiences. A story is the conceptual representation of what is told. Drawing on the work of Chatham and others (e.g. [17]) a story has central actors (which may for example be a person or a group), physical objects that appear in the story, and a theme. A story also contains a number of events. Any event may itself be composed of sub-events. An event has actors, a location, a specification of the time at which it happened, and additional properties depending on the type of event.
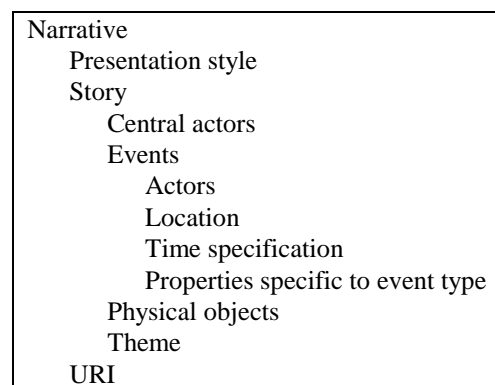
```
Narrative
    Presentation style
    Story
        Central actors
        Events
            Actors
            Location
            Time specification
            Properties specific to event type
        Physical objects
        Theme
    URI
```

Figure 2. The overall conceptual structure of narratives and stories in Story Fountain.

Exploration facilities

The conceptual structure provides the basis for the exploration facilities of Story Fountain. Story Fountain provides support for six exploration facilities. These are shown in table 1. The selection and design of these six functions were guided by a logical analysis of what was required to explore the story-base, what could be built above a conceptual representation of stories and narrative, and an investigation of narrative analysis techniques such as Story Network Analysis [2].

All of these six facilities are provided directly within the Story Fountain interface used by the Bletchley Park guides community. The Story Understanding (SU) facility provides a view of the conceptual structure of a story, modified for readability, in terms of its central characters, events, physical objects and themes. The Concept Understanding (CU) facility collects together all stories that contain a selected concept, whether that be a person, group, physical object, location or time period, and presents each story with a description of how it is related to the concept.

Concept Comparison (CCm) selects stories related to multiple concepts so that they can be compared and contrasted in a single representation. Stories can be automatically ordered to further support concept comparison, by aligning stories that have the selected concepts as values of the same property. For example, a Guide may compare and contrast stories concerned with Colossus (the world's first programmable computer developed at Bletchley Park) with the Bombe (the decryption machine designed and most widely used at Bletchley Park). A story containing an event describing that Colossus was designed by Tommy Flowers would be aligned with a story describing that the Bombe was designed by Alan Turing. Similarly, a story containing an event stating that Colossus was used to decrypt messages from the Lorenz encryption machine would be aligned with a story stating that the Bombe was used to decrypt messages from the Enigma encryption machine.

Name	Description
Story understanding (SU)	Understanding stories in terms of the concepts they contain and their relationship to them.
Concept understanding (CU)	Understanding concepts in terms of the how they are used in stories and in what context.
Concept comparison (CCm)	As in CU but comparing across multiple concepts.
Concept connecting (CCn)	Connecting concepts, finding out how they can be related through the stories.
Story mapping (SM)	Mapping story relationships, such as who lived where.
Event mapping (EM)	As in SM but on the event level.

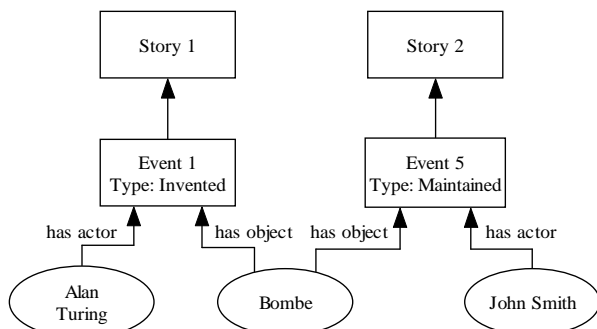**Table 1.** The six exploration facilities provided in Story Fountain.

Figure 3. A schematic illustration of concepts connected by stories.

Concept Connection (CCn) automatically provides pathways between concepts via the events contained in the stories. This facility, if for example used to connect the people Alan Turing and John Smith could provide the connection illustrated in figure 3. Event 1, contained in Story 1, describes Alan Turing's invention of the Bombe. Event 5, contained in Story 2, describes John Smith's maintenance of the Bombe machines. Therefore, two events in two different stories were required to connect these people. The output comprises a list of stories and describes how the story events are used to make connections between the selected concepts.

The Story Mapping (SM) facility provides a structure related to Story Network Analysis (described earlier). However, rather than just being available to map stories according to characters, Story Mapping can be applied to any story property represented in the knowledge base. For example, stories can be mapped in terms of the physical objects they describe. Event Mapping (EM) works in a similar way but with the properties of events rather than stories (such as event location).

Exploration processes
So far in this section we have described six exploration facilities provided by Story Fountain. Working in collaboration with Bletchley Park we have identified the types of open-ended research questions that they investigate using stories as the primary source material (e.g. first hand accounts and interviews).

These questions were further identified as being characteristic of the following exploration processes: accumulation, association and induction. Accumulation describes the aggregation of information of a particular type across a set of source materials. Association is the identification of contingencies between the events, objects and characters described in source materials. Induction refers to the analogical use of source materials to entertain unknown or hypothetical situations. These three processes encompass a broad range of the open-ended questions that can be pursued when conducting research with stories.

A fourth category, termed information gathering, was used to describe the type of exploration process that we can envisage being performed by visitors at the Park with a simplified interface to Story Fountain. Table 2 outlines how the four processes are supported by the six exploration facilities. Each exploration process is accompanied by an example question. If an abbreviation for an exploration facility is shown in brackets in the third column, this indicates that it is a supporting rather than primary information source for that exploration process.

An associated methodology is being developed for how these tools should be used to explore high-level questions. This will focus on how a question can be broken down into sub-questions whose answer can be sought by using the provided exploration facilities.

User	Exploration process	Exploration facilities
BP Guide	Accumulation: What was life like for wartime workers?	CU, SM, EM, (SU)
	Association: How did Bletchley Park influence the development of computers?	CU, CCm, SM, EM, (SU)
	Induction: Could Bletchley Park have decoded messages from the Type-X encryption machine?	CU, CCn, SM, EM, (SU)
Visitor	Information gathering: I would like to know more about Tommy Flowers and Alan Turing	CCm, CCn

Table 2. The four identified exploration process and the facilities available to support them.

USER SCENARIOS

The following two user scenarios, though fictional, illustrate how Story Fountain is used in practice to support the pursuit of open-ended, high-level questions using a corpus of digital stories. The first scenario describes how Story Fountain is used to support the research activities of a tour guide. The second scenario describes how a visitor to the Park is later provided with information relating to the exhibits which they found interesting during their visit.

Jane the Guide

Jane is a retired nurse and has lived in Bletchley for most of her life. Through her interest in local history she got involved as a volunteer tour guide at Bletchley Park shortly after she retired. A common question she is asked on her tours by visitors to the Park is 'what was life like for wartime workers?' Therefore, Jane periodically seeks out new information regarding wartime life. As part of her research she uses Story Fountain (see figure 5). Here she can get support with understanding stories, the concepts described in the stories, comparing concepts, connecting concepts, and mapping the story space according to the properties of the stories it contains.

As directed in the training Jane received in the use of Story Fountain Jane decides to break her open-ended research question into specific sub-questions. First she decides to find out where the wartime workers were billeted (i.e. housed). To do this she selects the billeting locations option from the 'Map events by' menu. After clicking on the 'GO' button she receives a list of stories and an overview of the events they describe for each billeting location (see figure 6).

From this list Jane can click on the links provided to view a story, the conceptual structure of a story, the events and

people described in a story. Jane decides to find out more about a specific billeting location, she clicks on the label for a village called Woburn Abbey to look at the stories about billeting in that village (see figure 7).
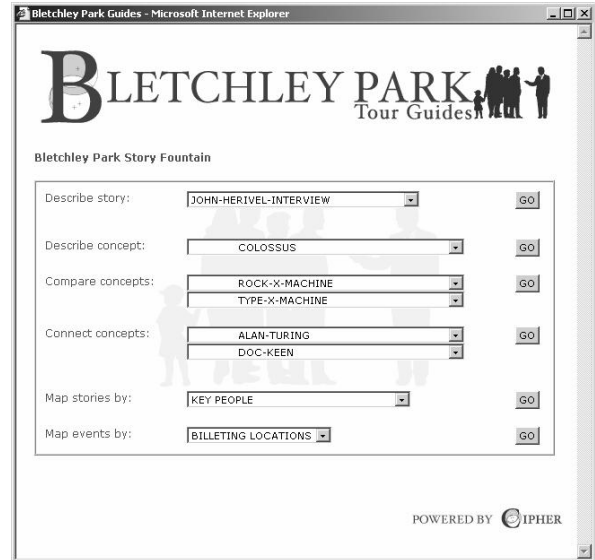

Figure 5. Home page for the Bletchley Park Guide version of Story Fountain.

Figure 6. An events map of billeting locations.

She wonders how a person's status within the Park affected their billeting location and conditions. Therefore, she selects one of the stories about life in Woburn Abbey and finds that the person described in the story worked as a Bombe operator (see figure 8).

She decides to compare this with the experiences of other Bombe operators. To do this, Jane clicks on the working role described, i.e. bombe-operator, in order to see which other stories describe people carrying-out this role (see figure 9).

Figure 7. A concept description of Woburn Abbey Village.

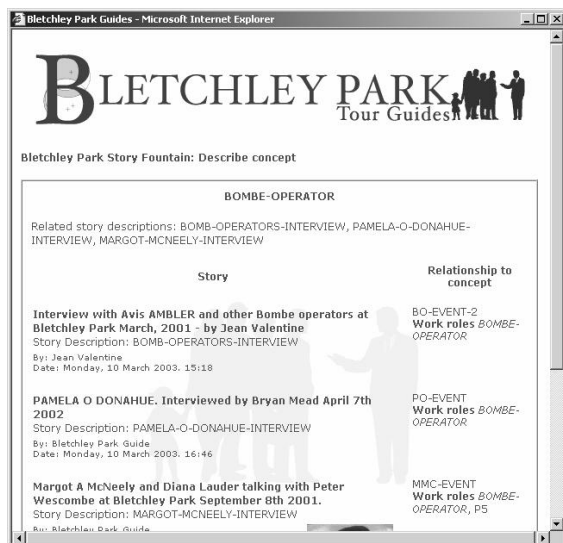**Figure 8.** A story description of an interview with Margot McNeely ad Diana Lauder

Figure 9. A concept description of a Bombe operator.

Jane finds several other stories that discuss life as a Bletchley Park Bombe operator. Reading these stories Jane extends her own knowledge of Bletchley Park and decides to integrate some of the personal experiences she has read about into the tours that she provides.

Adrian the Visitor

Adrian lives in Buckinghamshire with his wife and their two children. Adrian and his family decide go to Bletchley Park for a day out. When they arrive at the Park they are told that if there are any exhibits which they would like further information on they can send a text message identifying the exhibits of interest to the Park and later access a website containing further information when they get home.

During their visit the children take a particular interest in people working in cryptography, so they decide to send a text message requesting further information on two people they read about at the Park: John Tiltman and John Herivel. When the family arrives home they go to the Bletchley Park web site and are asked to enter their mobile phone number in order to view stories related to their interests (see figure 10).

Figure 10. The sign in page of the Bletchley Park visitor version of Story Fountain.

When they type in their phone number and press the 'GO' button, the message they sent is accessed and a set of Story Fountain queries are constructed using their chosen set of concepts. The results of the queries are then presented in a form tailored for the visitors to the Park (see figure 11). Although the same system is used as in the previous guide research scenario, the interface in this case does not allow exploration via concepts, only direct links to other stories. The underlying knowledge model is hidden from the user.

The resulting output page shows a list of the stories associated with each concept, and a story pathway showing how the two concepts can be linked through connections in the story-base. In this case the family finds out that John Tiltman was interviewed for his post at Bletchley Park at Cambridge University and so was John

Herivel. From this starting point the family is able to read more about the people, places and machines used at the time by following the numerous links to other parts of the story-base. The family decides to go to the Imperial War Museum and also to return to Bletchley Park to find out more about the people and places involved in decrypting ciphers.

Figure 11. A results page showing stories about John Tiltman (upper left) and John Herivel (upper right), and a story path connecting the two (lower half).

ARCHITECTURE

As noted in the introduction, Story Fountain uses ontologies of the domain and the heritage resources plus a reasoning engine for the selection and organization of resources. Ontologies used include the CIPHER Conceptual Reference Model Ontology (based on the ISO CIDOC CRM [5]), a domain ontology for the application, and a story and narrative ontology. These ontologies have been applied in order to produce a knowledge base for the Bletchley Park Forum. Within the architecture of the Forum input and presentation modules provide interfaces appropriate for specific audiences (see figure 12).
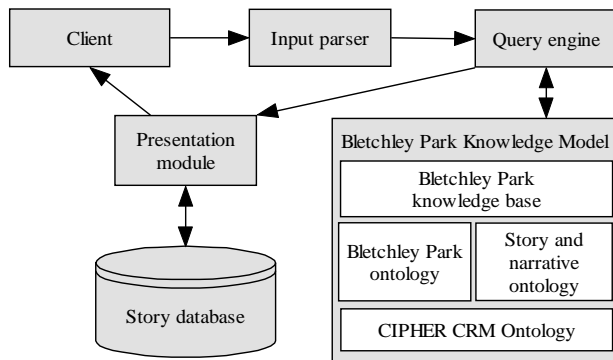

Figure 12. The system architecture of Story Fountain.

A user accessing Story Fountain via a client device (such as a web browser) can input a query which is passed to the query engine. The query engine interrogates the knowledge model and retrieves a set of knowledge resources. This set of resources are then passed to the presentation module which accesses the story database in order to retrieve additional story content which is formatted and presented to the user on their client device.

The ontologies were developed using the Apollo ontology development environment [1]. The resulting knowledge model is implemented in OCML [12] and runs on a Lisp server, which is accessible through an Apache Web Server using the 'mod_lisp' module [11]. MySQL is used to hold the story database and is accessible via an ODBC database connection. Python is the main scripting language used to produce the personalised input and presentation interfaces.

DISCUSSION & FUTURE WORK

Story Fountain has proved to be a useful tool for supporting research activities, it is currently used by the tour guides working at Bletchley Park and the use of the mobile phone application is to be trialed with visitors early in 2004.

The ongoing work with Story Fountain in the CIPHER project will focus on developing support for defining pathways across stories, actively supporting the reduction of a user's high-level research question into sub-questions suitable for exploration with Story Fountain, and the further development of Story Fountain as a generic tool for story research.

Currently the raw knowledge base is used to make connections between concepts by navigating across stories with common people, locations or objects. However, more useful goal-specific pathways may be defined by the user given a suitable interface. In this way a personal rule set may be used to navigate the knowledge model, such as; X could have passed knowledge to Z if X met Y at time t1, and Y met Z at time t2, and t1 occurred before t2.

Helping users to identify suitable sub-questions in order to explore a high-level research question may be supported in the future with the use of a research wizard. How to decompose problems and interpret data are key transferable skills useful for a wide range of tasks. A question and answer wizard may be a useful support for identifying the relevant concepts, formulating hypotheses, and comparing and contrasting data.

Developing a generic story research tool is a challenging task as it requires extensive support for customization and personalization. Ideally a user should be able to configure Story Fountain using a form-style interface, to identify the look and feel of the presentation, as well the location and format of the story-base and knowledge model. The adoption of such a tool would also require extensive support for the analysis of the domain's exploration

process and guidance on how the facilities provided by Story Fountain can support a broad range of exploration processes.

SUMMARY

In summary, this paper has presented the design, development and use of Story Fountain, a community support tool for query-driven story-base exploration. Story Fountain provides facilities to aid users engaged in the exploration processes of: accumulation, induction, association, and information gathering. This is achieved by integrating a powerful query engine with a comprehensive knowledge model. The knowledge model used here was built using the ISO CIDOC Conceptual Reference Model, along with our own domain, and story and narrative ontologies. The presentation of the query input and results are tailored to suit the user. This system is being used by the tour guides at Bletchley Park, further development and evaluation is under way.

REFERENCES

1. Apollo - http://apollo.open.ac.uk

2. Boje, D.M. *Narrative Methods for Organizational and Communication Research*. Sage, London. (2001).

3. Brusilovsky, P. and Maybury, M.T. From adaptive hypermedia to adaptive Web. *Communications of the ACM 45*, 5, (May 2002), 31-33. Special Issue on the Adaptive Web, P. Brusilovsky and M.T. Maybury (Eds.).

4. Chatman, S. *Story and Discourse: Narrative Structure in Fiction and Film.* Cornell University Press, New York. (1980).

5. CIDOC Conceptual Reference Model. Available at http://zeus.ics.forth.gr/cidoc/index.html

6. Collins, T.D., Mulholland, P. and Zdrahal, Z. Methodology and Tools to Support Storytelling in Cultural Heritage Forums, in *DEXA Workshop on Presenting and Exploring Heritage on the Web*, (Prague, Czech Republic, September 2003).

7. Dumais, S., Cutrell, E., and Chen, H. Optimizing search by showing results in context, in *Human Factors in Computing Systems (CHI '2001)*, (March 31-April 4, 2001, Seattle, WA, USA). 277-284.

8. Fellbaum C. (Ed.) WordNet: An Electronic Lexical Database. MIT Press. (1998).

9. Hein, G. Constructivist Learning Theory, in *The museum and the needs of people: Proceedings of ICOM-CECA Annual Conference*, (Jerusalem, Israel, 15–22 October 1991), 1992 ICOM-Israel, 89–94.

10. Hochheiser, H., and Shneiderman, B. Using interactive visualizations of WWW log data to characterize access patterns and inform site design. *Journal of the American Society of Information Science 52*, 4 (2001), 331-343.

11. Mod_Lisp - http://www.fractalconcept.com/

12. Motta, E. *Reusable components for knowledge modeling: Case studies in parametric design problem solving*, IOS, Amsterdam, The Netherlands, (1999).

13. Mulholland, P., Zdrahal, Z., and Collins, T.D. "CIPHER: Enabling Communities of Interest to Promote Heritage of European Regions", *Cultivate Interactive 8*, (15th November 2002). Available at http://www.cultivate-int.org/issue8/cipher

14. Muslea, I., Minton, S., and Knoblock, C.A. Hierarchical Wrapper Induction for Semistructured Information Sources. *Autonomous Agents and Multi-Agent Systems 4*, 1/2 (2001), 93-114.

15. Paley, N. *Finding Art's Place: Experiments in contemporary education and culture*, Routledge, London, UK. (1995).

16. Scanlon, E., Whitelegg, E., and Yates, S. (Eds) *Communicating Science: Contexts and Channels Reader 2*. Routledge, London, UK. (1999).

17. Schank, R.C. *Tell Me A Story: A New Look at Real and Artificial Memory*. Macmillan, New York. (1990).

18. Schneiderman, B. *Dynamic queries, starfield displays, and the path to Spotfire*. (February, 1999) Available at http://www.cs.umd.edu/hcil/spotfire/

19. Shipman III, F.M., Furuta, R., Brenner, D., Chung, C.C., and Hsieh, H. Using Paths in the Classroom: Experiences and Adaptations, in *UK Conference on Hypertext* (UK, 1998), 267-270.

20. Stuer, P., Meersman, R., and De Bruyne, S. The HyperMuseum Theme Generator System: Ontology-based Internet support for the active use of digital museum data for teaching and presentation, in D. Bearman & J. Trant (Eds) *Museums and the web 2001: Selected Papers.*, Archives & Museum Informatics, 2001.