

Expanding a Humanities Digital Library: Musical References in Cervantes' Works

Manas Singh[§], Richard Furuta[§], Eduardo Urbina[¶], Neal Audenaert[§], Jie Deng[§],
Carlos Monroy[§]

Center for the Study of Digital Libraries*
Texas A&M University
College Station, TX, 77843-3112
cervantes@cSDL.tamu.edu

Abstract. Digital libraries focused on developing humanities resources for both scholarly and popular audiences face the challenge of bringing together digital resources built by scholars from different disciplines and subsequently integrating and presenting them. This challenge becomes more acute as libraries grow, both in terms of size and organizational complexity, making the traditional humanities practice of intensive, manual annotation and markup infeasible. In this paper we describe an approach we have taken in adding a music collection to the *Cervantes Project*. We use metadata and the organization of the various documents in the collection to facilitate automatic integration of new documents—establishing connection from existing resources to new documents as well as from the new documents to existing material.

1 Introduction

As a digital library grows in terms of both size and organizational complexity, the challenge of understanding and navigating the library's collections increases dramatically. This is particularly acute in scenarios (e.g., scholarly research) in which readers need and expect to be able to survey all resources related to a topic of interest. While large collections with a rich variety of media and document sources make valuable information available to readers, it is imperative to pair these collections with tools and information organization strategies that enable and encourage readers to develop sophisticated reading strategies in order to fully realize their potential [8]. Traditional editorial approaches have focused on detailed hand editing—carefully reading and annotating every line on every page with the goal of producing a completed, authoritative edition. Often, such approaches are infeasible in a digital library environment. The sheer magnitude of many digital collections (e.g., the Gutenberg Project [17], the *Christian Classics Ethereal Library* [26], the *Making of America* [8][22]) make detailed hand editing unaffordably labor intensive, while the very nature of the project often conflicts with the traditional goal of producing a final,

* Authors' academic affiliations: [§]Department of Computer Science, Texas A&M University; [¶]Department of Hispanic Studies, Texas A&M University

fixed edition. Previously, we have described the multifaceted nature of humanities collections focused on a single author and argued that these projects will require automatic integration of many types of documents, drawn from many sources, compiled by many independent scholars, in support of many audiences [2]. Such collections are continuously evolving. As each new artifact is added to the collection, it needs to be linked to existing resources and the existing resources need to be updated to refer to the new artifact, where appropriate. Constructing these collections will require new tools and perspective on the practice of scholarly editing [11]. One such tool class is that supporting automatic discovery and interlinking of related resources.

The *Cervantes Project* has been focused during the last ten years on developing online resources on the life and works of Miguel de Cervantes Saavedra (1547 – 1616), the author of *Don Quixote* [32], and thus has proven to be a rich environment for exploring these challenges. Given its canonical status within the corpus of Hispanic literature and its iconic position in Hispanic culture, the *Quixote* has received a tremendous amount of attention from a variety of humanities disciplines, each bringing its own unique questions and approaches. Within the broad scope of this project, individual researchers have made a variety of contributions, each centered on narrowly scoped research questions. Currently, work in the project can be grouped into six sub-projects: bibliographic information, textual studies, historical research, music, *ex libris*, and textual iconography. Together, these contributions span the scope of Cervantes' life and works and their impact on society.

In this paper, we describe the approach that we have taken in connection with the presence and influence of music in Cervantes' works. The data for this project was collected by Dr. Juan José Pastor as part of his dissertation work investigating Cervantes' interaction with the musical culture of his time and the subsequent musical interpretations of his works [21]. Pastor's collection is organized in five main categories (instruments, songs, dances, composers, and bibliographical records) and contains excerpts from Cervantes' writings, historical and biographical information, technical descriptions, images, audio files, and playable scores from songs. Although Pastor has completed his dissertation, the collection is still growing, as new scores, images, and documents are located. For example, a recent addition, produced in conjunction with the 400th anniversary of the publication of the *Quixote*, is a professionally-produced recording of 22 of the songs referred to by Cervantes [13].

The music sub-project reflects many aspects of the complexity of the *Cervantes Project* as a whole, and thus provides an excellent testbed for developing tools and strategies for integrating an evolving collection of diverse artifacts for multiple audiences. A key challenge has been determining how to build an interface that effectively integrates the various components, in a manner that supports the reader's understanding of the implicit and explicit relationships between items in the collection. In particular, since the collection is growing with Pastor's ongoing research, it was necessary that the interface be designed so that new resources could be easily added and the connections between new and old resources generated automatically. To address this challenge we have developed an automatic linking system that establishes relationships between resources based on the structural organization of the collection and various metadata fields associated with individual documents. An editor's interface allows users an easy way to add new resources to the

collection and to specify the minimal set of metadata required to support link generation. Further, a reader's interface is provided that identifies references within texts to other items in the collection and dynamically generates navigational links.

2 Background

Developing a system to integrate resources within the collection required attention to three basic questions: What types of reader (and writer/editor) interactions are to be supported? What types of information and connections are to be identified? How will that information be identified and presented to readers? A brief survey of related projects will help to set the context for the design decisions we have made in these areas.

The *Perseus Project* has developed a number of sophisticated strategies for automatically generating links in the context of cultural heritage collections. Our work has been heavily influenced by their use of dense navigational linking both to support readers exploring subjects with which they are unfamiliar and to encourage readers more closely acquainted with a subject to more fully explore and develop their own interpretive perspectives. Early work focused on developing language based tools to assist readers of their extensive Greek and Latin collections. These tools linked words to grammatical analysis, dictionaries and other linguistic support tools, helping a wider audience understand and appreciate them. More recently, they have focused on applying some of the techniques and technologies developed for their Classical collection to a variety of other, more recent data sets including American Civil War and London collections. This work has focused on identifying names, places, and dates to provide automatically generated links to supplementary information and to develop geospatial representations of the collection's content. They have had good results from a layered approach using a combination of *a priori* knowledge of semi-structured documents (e.g., of the *British Directory of National Biography* and *London Past and Present*), pattern recognition, name entity retrieval, and gazetteers to identify and disambiguate references to people, places, and events.

A key technology for supporting this type of integration between resources within a collection is the use of name authority services. The SCALE Project (Services for a Customizable Authority Linking Environment) is developing automatic linking services that bind key words and phrases to supplementary information and infrastructure to support automatic linking for collections within the National Science Digital Library [24]. This collaborative effort between Tufts University and Johns Hopkins University builds on the tools and techniques developed in the Perseus Project in order to better utilize the authority controlled name lists, thesauri, glossaries, encyclopedias, subject hierarchies and object catalogs traditionally employed in library sciences in a digital environment.

As an alternative to authority lists, the Digital Library Service Integration (DLSI) project uses lexical analysis and document structure to identify anchors for key terms within a document [6]. Once the anchors are identified, links are automatically generated to available services based on the type of anchor and the specified rules.



Figure 1: Related Links and a Sample Image for the *Sonaja* Instrument

For example, if a term is a proper noun it can be linked to glossaries and thesauri to provide related information.

Also of relevance is the long history in the hypertext research community of link finding and of link structures that are more than simple source to destination connections. Early work in link finding includes Bernstein’s Link Apprentice [5] and Salton’s demonstration of applications [29] of his Smart system’s vector-space model [28]. Link models related to our work include those that are multi-tailed, for example MHMT [20] and that represented in the Dexter model [18].

3 Interface and Usage Scenario

Within the context of the Cervantes music collection, we have chosen to focus on identifying interrelationships between the structured items in our collection in order to provide automatic support for the editorial process rather than relying on authority lists or linguistic features to connect elements of the collection to externally supplied information sources (such support for this could be added later, if warranted). We have divided the resources in our collection into categories of structured information (e.g., instruments, songs, composers). Each category contains a set of items (e.g., a particular song or composer). Each item is in turn represented by a structured set of documents. How the documents for any given item are structured is determined by the

category it is a member of. For example, *arpa* (a harp) is an item within the instruments category. This instrument (like all other instruments) may have one or more of each of the following types of documents associated with it: introductory articles, images, audio recordings, historical descriptions, bibliographic references, links to online resources, and excerpts from the texts of Cervantes that refer to an *arpa*.

Each item is identified by its name and by a list of aliases. Our system identifies the references to these terms in all of the resources located elsewhere in the collection, either as direct references or within the metadata fields of non-textual documents. At present, the matching algorithm is a simple match between the longest-length term string found at the target. Once identified, the references are linked to the item.

The presentation of information to the reader uses a custom representation of links. This is because of the complexity of the object linked to—a complexity that reflects the multiple user communities that we expect will make use of the collection. Moreover, the collection provides multiple roots that reflect different reader specializations.

In developing the Cervantes music collection we have focused our design on meeting the needs of two primary communities of readers. One group is composed of Cervantes scholars and music historians interested in research about Cervantes' works and music. The second group is composed of non-specialists interested in gaining access to information they are unfamiliar with. For both the specialist and the non-specialist, the collection provides two major focal points, or roots, for access. For example, a reader might approach the music collection from the texts of Cervantes (which themselves compose a distinct collection), asking how a particular passage reflects Cervantes' understanding of contemporary musical trends or in order to better understand what, for example, an *albugue* looks and sounds like.¹ Another reader might begin by considering a particular composition that alludes to Cervantes and ask how this particular piece reflects (or is distinct from) other popular interpretations of the *Quixote*. Similarly, a non-expert might find his understanding of a particular opera enhanced by learning more about an obscure reference to one of Cervantes' works. In this way the linkages generated between these two distinct but related collections allow readers access to a rich and diverse body of resources from multiple perspectives to achieve a variety of goals. We refer to collections that exhibit this type of structure as being multi-rooted. Natural roots for the music collection include compositions (e.g., songs and dances), composers, instruments, and the writings of Cervantes. In the remainder of this section we present several brief reader interaction scenarios to help illustrate the design of the system from a reader's perspective. In the following section we present an overview of the technical design and implementation of the link generation system and the interface.

¹“What are albugues?” asked Sancho, “for I never in my life heard tell of them or saw them.” “Albugues,” said Don Quixote, “are brass plates like candlesticks that struck against one another on the hollow side make a noise which, if not very pleasing or harmonious, is not disagreeable and accords very well with the rude notes of the bagpipe and tabor. [Chapter 65, Part 2, *Don Quixote*]



Figure 2: Learning more about the arpa.

In the first scenario, a native, modern Spanish speaker is reading a less well-known text of Cervantes, *Viaje del Parnaso*, and encounters a reference to an instrument she is unfamiliar with, the *sonaja*. Curious, she clicks on the link and a drop-down menu appears displaying links to the various types of documents present in the collection. She elects to view the ‘sample image,’ resulting in the display shown in Figure 1. The image sparks her curiosity and she decides to see what it sounds like by clicking on the ‘sample audio’ link. What is this, who would use it, and why? To find out more, she clicks to read the introductory text and finds a list of definitions where she learns that it is a small rustic instrument that was used in the villages by beating it against the palm of the hands. Interestingly, the Egyptians used it in the celebrations and sacrifices to the god Isis. Having learned what she wanted to know, she returns to reading *Viaje del Parnaso*.

In the second scenario, a music historian with relatively little familiarity with *Don Quixote* or the other works of Cervantes is interested in exploring how string instruments were used and their societal reception. On a hunch, he decides to see how societal views of the harp and other instruments might be reflected in the works of Cervantes. Browsing the collection, he navigates to the section for the harp and peruses the texts of Cervantes that refer to the harp (Figure 2). After surveying that information, he explores some of the other instruments in order to get a broader perspective on how Cervantes tends to discuss and incorporate musical instruments in his writings. He finds a couple of passages that help to illustrate the ideas he has been developing, and makes a note of them to refer to later.

In the final scenario, an editor is working with the collection, adding the historical documents to the song, “Mira Nero de Tarpeya.” As shown in Figure 3, he browses to the list of composers and notices that, while there is a link to Mateo Flecha, there is

no information provided for Francisco Fernández Palero. He quickly navigates to the “composers” category, adds Palero as a new composer (Figure 4), and writes a short description of him and his relevance to classical music. The system recognizes the new composer and updates its generated links accordingly. Currently, since only minimal information is present, these links refer only to the newly written introductory text. A few weeks later, the editor returns to the collection after finding images, lists of songs written, and historical descriptions. He adds these through forms similar to the one he used to add Palero. Links to these new resources are now added to the drop down menu associated with references to Palero. In this way, the editor is able to focus on his area of expertise in finding and gathering new information that will enhance the scholarly content of the collection, removing the burden of manually creating links from all the existing documents to the newly added composer.

Categories	Mira Nero de Tarpeya		
Instrument	Explanation	Music Note	Audio
Song	Bibliography	Composer	Cervantes Text
Dance	Click here to add a new Composer entry		
Composer	List of Composer follows		
Bibliography	Mateo Flecha		
Musical Reception	el Viejo, Las ensaladas de Flecha, Maestro de capilla que fue de las Serenísimas Infantas de Castilla. Recopiladas por el F. Matheo Flecha, su sobrino..., Praga, Jorge Negrito, 1581. (Barcelona, Biblioteca de Catalunya, Ms. 851; vid. también Ms. 588/1 y 588/2).		
Virtual Tour	Francisco Fernández Palero		
Full Text	Libro de Cifra Nueva. Para tecla, harpa y vihuela, Alcalá, Iona de Brocar, 1557 (Madrid, Biblioteca Nacional de España, R. 598)		
Log Out			

Figure 3: Browsing a Song in the Editor’s Interface

Categories	Composer		
Instrument	Add a new Composer		
Song	Name :	<input type="text" value="Francisco Fernández Palero"/>	
Dance	Alternative Name :	<input type="text" value="Palero"/>	
Composer	Alternative Name :	<input type="text"/>	
Bibliography	Alternative Name :	<input type="text"/>	
Musical Reception	Alternative Name :	<input type="text"/>	
	<input type="button" value="Submit"/>		<input type="button" value="Cancel"/>

Figure 4: Adding the Composer Francisco Fernández Palero

4 Organization of the Digital Library

Information in the collection is organized as hierarchical groups. At the highest level, materials are grouped into eight categories:

- 1 Instruments: information pertaining to the different musical instruments that have been referred to by Cervantes in his works.
- 2 Songs: information regarding the different songs that have influenced Cervantes.
- 3 Dances: resources related to the dances that have been referred to in Cervantes' texts.
- 4 Composers: the composers who have influenced Cervantes and his work.
- 5 Bibliography: bibliographical entries related to instruments, songs, and dances that have been referred to in Cervantes' texts.
- 6 Musical Reception: bibliographical entries about musical compositions that have been influenced by Cervantes or refer to his works.
- 7 Cervantes Text: full texts of Cervantes' works.
- 8 Virtual Tour: links to virtual paths, constructed and hosted using Walden's Paths [30]. This allows the information to be grouped and presented in different manners, catering to the interests of diverse scholars, thus opening up the digital library to unique interpretive perspectives.

Most categories are further subdivided into items. An item defines a unique logical entity, as appropriate for the type of category. For example, the category "Instruments" contains items such as *arpa* and *guitarra*. Similarly, each composer would be represented as an item in the respective category as would each dance and each song. The item is identified by its name, perhaps including aliases (e.g., variant forms of its name).

Artifacts associated with each item are further categorized into different topics like image, audio, and text. The topics under an item depend on the category to which the item belongs to. For example, an item under category "Instruments" will have topics like introduction, audio, image, text, and bibliography but an item under the category "Composer" will have topics like life, image, work, and bibliography.

An artifact (e.g., an individual picture; a single essay) is the "atomic" unit in the collection. Thus artifacts are grouped under topics, which in turn are grouped into items, which in turn are grouped into categories. A unique item identifier identifies each item in the digital library. Additionally, each artifact placed under an item is assigned a sub-item identifier that is unique among all the artifacts under that item. Thus all the artifacts, including texts, audio files, images, musical scores, etc., are uniquely identified by the combination of item identifier and sub-item identifier.

5 Interlinking

The process of creating interlinks and presenting the related links can be broadly classified into four major steps. The first is maintaining the list of item names for which information exists in the digital library. The second is a batch job, which identifies the reference of these terms in all the texts present in the digital library. The

third step is a run time process, which, while displaying a text, embeds the terms that need to be linked with a hyperlink placeholder (i.e., hyperlink without any specific target). This step uses the data from the batch job to identify the terms that should be presented with the hyperlink for any text. The final step generates the actual related links for a term and is invoked only when the user clicks on a hyperlink placeholder. A description of these steps follows.

Maintaining the keyword list: In order for the system to provide related links, it should be able to identify the terms for which information exists in the digital library. This is achieved by maintaining a keyword list. To identify the variation in names a synonym list is also maintained. The system depends on the user to provide a list of synonyms for the item being added. This may include alternate names for the item or just variations in the spelling of the item name.

When a new item is added to the digital library its name or title is added to the keyword list and its aliases to the synonym list. In the following sections the keyword and synonym lists will be referred to collectively as keywords.

Document keyword mapping batch job: The document keyword mapping is created by indexing all the texts using Lucene and finding the references of each term in the keyword list among all the texts. This is done offline using a batch process. This also populates a document keyword map that maps each document to all the keywords it refers.

Runtime display of texts with hyperlink placeholders: While displaying a text the system uses the document keyword map to identify the keywords from the keyword list that are present in the text. Once the list of keywords present in the text is known, their occurrences in the text are identified and are embedded with hyperlink placeholders. In essence, instances of *keyword* in the source are replaced by,

```
<a href="javascript:nop" class="cerhyperlink"> keyword </a>
```

which invokes the appropriate display function when selected.

Display of composite links: The related links display is generated when the user clicks on a keyword's hyperlink placeholder. The click event is intercepted by a client side JavaScript function that parses the hyperlink statement and retrieves the actual keyword, sending the keyword to the server

When the request is received at the server, the keyword is retrieved from the request parameters and the metadata repository is used to find all the artifacts related to the keyword. Using these related artifacts, the distinct list of topics to which they belong is identified and links to these topics are generated. For example, if the item has some related image resources then a link to view the images is added to the related links list. Furthermore, the format of these artifacts is also noted. If they are of formats like image or audio then a link to a sample image or audio also is added to the related links list. This sample audio or image is displayed in a new page right on top of the text. This allows the user to view a sample image or listen to a sample audio clip without leaving the text.

The response from the server is received by the client as an XML document object, which is parsed using JavaScript to obtain the related links. The related links are displayed in a tooltip just below the keyword clicked by the user. Cascading style sheets are used to control the look and feel of the tooltip.

6. Discussion and Future Work

This work is one of three major directions we are pursuing to better understand how complex, highly interdisciplinary humanities collections can be designed to enable tight integration resulting in a single, multi-rooted collection. Our focus in the Cervantes music collection has been on leveraging structural information captured as a natural part of the collection building process. Using this information, we are able both to identify link anchors (references to items in the collections) and resources to connect them to. The resulting navigational hypermedia archive enhances the reader's ability to access and interact with the collection as a whole. We would like to expand this approach by more formally investigating the types of structures that can be included in system such as this and the types of automatic linking strategies that each of these structures might support. For example, how might hierarchically structured categories be incorporated? How might that affect the types of interlinkages that can be established? Such an investigation will help us better understand what additional structures are needed to scale this approach to incorporate the full breadth of resources included within the *Cervantes Project* and how it could be generalized to meet the needs of other humanities projects.

In addition to the structural approach to integrating resources presented here, we have also reported on work that uses a formal narrative and thematic taxonomy to provide an integrative framework [3], and also the use of a framework for identifying key features within documents [4]. More work is needed to bring these three directions together to form a unified approach and to understand how each contributes to the larger goal of single, multi-rooted collection.

As we are developing these ideas, we are becoming more aware of the need for a shift in the way we understand the editorial process. Traditional editorial work is focused on the development of a single, centered, completed work that is relatively fixed over time—a published edition. Despite the growing calls to shift from the book as the primary technology for developing scholarly editions to electronic media [21], the resulting editions bear much similarity to their ancestors. In particular, they retain the notion of a “completed” work that is developed to meet narrowly defined research objectives. They are typically created by a single editor or by the highly coordinated efforts of a group of authors working under the guidance of an editorial board. Such editions do not allow for the more complex types of informational needs that are required to support a more broadly defined humanities research agenda, such as that of the *Cervantes Project* [2]. This type of work is open ended and difficult to restrict to a closed set of scholarly perspectives—new research directions continually pop up, often initiated by people outside of the core project members. Its contributors are not the carefully orchestrated cadre of authors one might find in a scholarly encyclopedia (e.g., the Stanford Encyclopedia of Philosophy [31]), but rather are individual researchers pursuing their own unique research ideas (and making their own unique contributions). These researchers will often be uncooperative, if for no other reason than their divergent interests (the ethnomusicologist is not likely to be overly concerned about the work of the textual critic), yet their research may contribute significantly to the broader goals of such a digital library. This is not to suggest that traditional approaches are bad or should be abandoned, but rather to propose that we

need to creatively explore how to best employ digital technologies to empower humanities research.

References

- [1] "Ajax Tags," <http://ajaxtags.no-ip.info/>, Accessed 2nd Jan 2006.
- [2] N. Audenart, R. Furuta, E. Urbina, J. Deng, C. Monroy, R. Saenz, and D. Careaga, "Integrating diverse research in a digital library focused on a single author," *Proc. 9th European Conf. on Research and Advanced Technology for Digital Libraries*, Vienna, Austria, 2005.
- [3] N. Audenart, R. Furuta, E. Urbina, J. Deng, C. Monroy, R. Saenz, and D. Careaga, "Integrating collections at the Cervantes Project," *Proceedings of the 5th ACM/IEEE-CS joint conference on Digital Libraries*, Denver, USA, 2005.
- [4] N. Audenaert, R. Furuta, and E. Urbina, "A General Framework for Feature Identification," *Digital Humanities 2006*, to appear.
- [5] M. Bernstein, "An Apprentice that Discovers Hypertext Links," *Proceedings of the European Conference on Hypertext*, November 1990, pp. 212-223.
- [6] Miguel de Cervantes M (1998) *Don Quijote de la Mancha*, Francisco Rico, Director. Barcelona: Biblioteca Clásica, 2 vols; *Don Quixote* (English translation by Edith Grossman) New York: HarperCollins, 2003.
- [7] Xin Chen, Dong-ho Kim Kim, N. Nnadi, H. Shah, P. Shrivastava, M. Bieber, Il Im, and Yi-Fang Wu, "Digital Library Service Integration," *Proceedings of the 3rd ACM/IEEE-CS joint conference on Digital Libraries*, pp. 384-384, Houston, Texas, 2003.
- [8] Cornell University (2005) Making of America <http://moa.cit.cornell.edu/moa/> [accessed 8 Sept 2005]
- [9] G. Crane, (1998) The Perseus Project and beyond. D-Lib Magazine. <http://www.dlib.org/dlib/january98/01crane.html>
- [10] G. Crane (2000) Designing documents to enhance the performance of digital libraries: time, space, people and a digital library of London. D-Lib Magazine 6 (7/8). <http://www.dlib.org/dlib/july00/crane/07crane.html>
- [11] G. Crane, J.A. Rydberg-Cox, (2000) New technology and new roles: the need for "corpus editors." *Proc. 5th ACM conference on digital libraries*, San Antonio, TX, pp 252-253
- [12] G. Crane, E. David, A. Smith, and C. E. Wulfman, "Building a Hypertextual Digital Library in the Humanities: A Case Study on London," *Proc. 1st ACM/IEEE-CS joint conference on Digital libraries*, pp. 426-434, Roanoke, Virginia, United States, 2001.
- [13] Ensemble Durendal. *Por ásperos caminos. Nueva música cervantina*. Ediciones de la Universidad de Castilla-La Mancha, Cuenca, 2005. Text by J. J. Pastor and musical direction by S. Barcellona.
- [14] E. Frank, G. W. Paynter, I. H. Witten, C. Gutwin, and C. G. Nevill-Manning, "Domain-Specific Keyphrase Extraction," *Proc. Sixteenth International Joint Conf. Artificial Intelligence*, Morgan Kaufmann Publishers, San Francisco, CA, 1999.
- [15] R. Furuta, S. S. Kalasapur, R. Kochumman, E. Urbina, and R. Vivancos-Perez, "The Cervantes Project: Steps to a customizable and Interlinked On-Line Electronic Variorum

Edition Supporting Scholarship,” *Proc. 5th European Conf. Research and Advanced Technology for Digital Libraries*, Darmstadt, Germany, 2001.

- [16] G. Geisler, S. Giersch, D. McArthur, and M. McClelland, “Creating virtual collections in digital libraries: benefits and implementation issues,” *Proc. 2nd ACM/IEEE-CS joint conf. on Digital Libraries*, pp. 210-218, Portland, Oregon, USA, 2002.
- [17] Project Gutenberg Literary Archive Foundation (2005) Project Gutenberg <http://www.gutenberg.org/>. [accessed 9 Sept 2005]
- [18] F. Halasz and M. Schwartz. The Dexter Hypertext Reference Model. *Communications of the ACM*, 37(2), February 1994. pp. 30-39.
- [19] S. Jones and G. Paynter, “Topic based browsing within a digital library using keyphrases,” *Proc. fourth ACM conf. Digital libraries*, pp. 114-121, Berkeley, California, United States, 1999.
- [20] B. Ladd, M. Capps, D. Stotts, and R. Furuta. “Multi-head/Multi-tail Mosaic: Adding Parallel Automata Semantics to the Web,” *Proc. 4th WWW Conf.*, pp. 422-440, 1995
- [21] J. McGann, *The Rationale of Hypertext* <http://www.iath.virginia.edu/public/jjm2f/rationale.html>
- [22] University of Michigan (2005) Making of America. <http://www.hti.umich.edu/m/moagr/> [accessed 8 Sept 2005]
- [23] J. J. Pastor, “Música y literatura: la senda retórica. Hacia una nueva consideración de la música en Cervantes,” Doctoral Dissertation, Universidad de Castilla-La Mancha, 2005.
- [24] M. S. Patton and D. M. Mimno, “Services for a customizable Authority Linking Environment,” *Proc. 4th ACM/IEEE-CS joint conf. Digital Libraries*, pp. 420-420, Tuscon, AZ, 2004.
- [25] L.D. Paulson, “Building rich web applications with Ajax,” *Computer*, vol. 38, issue 10, pp. 14-17, Oct 2005.
- [26] H. Plantinga, coord. (2005) Christian Classics Ethereal Library, Calvin College, Grand Rapids, MI. <http://www.ccel.org/> [accessed 8 September 2005]
- [27] B. Pritchett, “KeyLinking: Dynamic Hypertext in a digital library,” *Proc. fifth ACM conf. on Digital Libraries*, pp. 242-243, San Antonio, Texas, United States, 2000.
- [28] G. Salton, *Automatic Text Processing: The Transformation, Analysis, and Retrieval of Information by Computer*, Addison-Wesley, Reading, MA, 1989.
- [29] G. Salton, J. Allan, C. Buckley and A. Singhal, “Automatic Analysis, Theme Generation, and Summarization of Machine-Readable Texts,” *Science*, Vol. 264, No. 5164, Jun. 1994, pp. 1421-1426.
- [30] F. Shipman, R. Furuta, D. Brenner, C. Chung, and H. Hsieh, “Guided Paths through Web-based Collections: Design, Experiences, and Adaptations,” *Journal of the American Society of Information Sciences (JASIS)*, 51(3), March 2000, pp. 260-272.
- [31] “Stanford Encyclopedia of Philosophy,” <http://plato.stanford.edu/>, [accessed March 7, 2006].
- [32] “The Cervantes Project,” E. Urbina, director. Center for the Study of Digital Libraries, Texas A&M University, [http://csdl.tamu.edu/cervantes.](http://csdl.tamu.edu/cervantes/), [accessed Nov 29 2005].