Textual Iconography of the *Quixote*; a Hypertextual Digital Archive at the Cervantes Project

Eduardo Urbina, Richard Furuta and Steven Smith
Texas A&M University

At present, there is no collection or catalogue, in print or online, covering in any comprehensive manner the textual iconography of the *Quixote*. Some attempts were made first in 1879 and later in 1905 to offer a representative sampling but the coverage is very limited in both cases, amounting in one case to 101 illustrations from 60 selected editions and in the other to 23 plates.\(^1\) Also in 1905 the *Iconografía de las ediciones del Quijote*, included 611 facsimile images but only of title pages. A more ambitious scholarly attempt was made in 1947, on the occasion of the 300th anniversary of Cervantes’ death, by Juan Givanel, who authored a *Historia gráfica de Cervantes y el Quijote* in which he includes and analyzes a total of 77 illustrations.\(^2\) In contrast, it is worth pointing out that one single edition (Paris, 1836) contains 800 engravings and drawings illustrating Cervantes’ story.\(^3\)

In the last few years the critical interest on the illustrations of the *Quixote* has been renewed as demonstrated by the publication of three major monographs by J. Hartau (Berlin, 1987), R. Paulson (Baltimore, 1998), and R. Schmidt (Montreal, 1999).\(^4\) These studies place the illustrations in new and diverse cultural, aesthetic and historical contexts, demonstrating their key critical value and role in the reception and interpretation of the novel, and make evident the urgent need to provide a more complete and accessible resource to the rich artistic tradition of the textual iconography of the *Quixote* in order to better understand its significant contribution to the editorial history and critical reception of Cervantes’ novel still largely unknown to readers and unexamined by critics.

Two obstacles have until now prevented the publication of a comprehensive collection or archive based on the textual iconography of the *Quixote*: 1) the rarity of and difficult access to the materials, and 2) the technical and financial difficulties in compiling and disseminating such an archive in print format. The advent of hypertext, digital libraries, and the Internet, among other technological factors of the information technology revolution, make the impossible dream of visualizing the *Quixote* a realizable goal. Nevertheless, in order to achieve such a goal there remain still considerable obstacles and challenges if the result is to be both effective and valuable as an educational tool and a research resource in the humanities.

The *Cervantes Project* (CP) will address the above described need through the creation of a fully accessible, searchable and documented electronic database and digital archive of all the illustrations that form the textual iconography of the *Quixote*, as permitted by copyright limitations, along with the necessary interfaces and visualization tools to allow for the kind of access and knowledge until now unavailable.\(^5\) We further envision the archive as a research depository to complement the textual and
A number of key factors and unique circumstances come together in our project: 1) the propitious event of the centenary and the renewed interest it brings to all aspects of Cervantes scholarship, 2) the obvious and urgent need to create electronic online access to rare collections and unique image resources until now impossible to achieve, 3) our existing collaborations with the Biblioteca Nacional of Spain and the University of Castilla-La Mancha, 4) the presence at Texas A&M University of a core and focused collection of illustrated editions of the Quixote, 5) the academic support and computer infrastructure provided by the Cushing Memorial Library and the Center for the Study of Digital Libraries, and 6) the scholarly experience and technological expertise developed over the last eight years by the members of the Cervantes Project in previous projects. We believe these advantageous circumstances represent a unique research and development opportunity to achieve the goals described.

In 2005 the world will celebrate the 400th anniversary of the publication of the Quixote, an occasion which promises to be not only a major cultural and political event, particularly in Spain and Spanish speaking countries, but a time both for celebration and reexamination, and one in which to appropriately frame the ongoing contributions of the Cervantes Project and the specific goal of this project: the creation of a hypertextual digital archive based on the textual iconography of the Quixote. This rich critical and artistic tradition remains largely unknown due in great part to the rare and inaccessible nature of the editions in which the thousands of illustrations that constitute the visual narrative and interpretation of Cervantes’ masterpiece have appeared. Although some sample illustrations are often reproduced in critical studies, catalogues and expositions, we unfortunately still lack knowledge of and access to the complete iconographic tradition associated with the publication of the Quixote as a major contributing element to both the canonization of the novel and to the iconic transformations of its principal character.

The main rare book collection supporting our project is the Cervantes Project Collection at the Cushing Memorial Library and Archives of Texas A&M University. In recent years, the Cervantes Project and the Cushing Memorial Library have acquired a large number of significant illustrated editions for the purpose of creating a collection specialized in illustrated editions of the Quixote. The Quixote textual iconography collection includes at present (May 2004) 250 illustrated editions, published since 1620. The collection comprises almost 600 volumes and is concentrated in 18th and 19th century English, French, and Spanish illustrated editions. We estimate the digital archive of the bibliographical electronic resources already present in the CP, as well as a unique digital variorum image collection able to extend the value of our Electronic variorum edition of the Quixote, a project funded by a grant from the National Science Foundation Information Technology Research program (2000-2003) and as a special project of the Digital Library Initiative (DL12). This comprehensive archive will allow worldwide electronic access to unique and rare textual and graphic resources by scholars, students and users in general interested in Cervantes’ work and on the impact and influence of his masterpiece through 400 years from several perspectives: textual, artistic, critical, bibliographical, and historical.
collection will eventually include upwards of 7,000 images and a fully searchable database complimented by innovative visualization tools.\textsuperscript{7}

The identification, acquisition and incorporation into our collection of the illustrated editions from which to extract and document the editorial history and textual iconography of the Quixote was initiated in 2000. As the Cervantes collection expands and new editions are acquired each year the digital archive will expand and achieve greater value and refinement. In addition, the Cushing Memorial Library of Texas A&M University will be providing metadata support, digitization coordination and oversight, management of the physical collection, cataloging expertise, access and acquisition oversight.

These developmental efforts and ongoing support ensure the comprehensiveness and future value of our collection and digital archive, as one of the preeminent rare book Cervantes collections in the country. On the technical side, the ongoing collaboration with the CSDL and our leading role and participation in the Humanities Informatics initiative started by TAMU in 2002 also ensures the availability of continued technical and infrastructure support. In addition, the collaboration recently established through the Cervantes Chair at the University of Castilla-La Mancha will also provide additional resources and personnel through faculty exchanges and support for graduate student research.\textsuperscript{8}

An important component of our initial work is the specification of a comprehensive taxonomy of the episodes, adventures, themes and characters in the Quixote. An initial prototype of the taxonomy designed in collaboration with Jesús Maestro. The taxonomy, representing the logical narrative structure of the work, will provide the addressing mechanism by which illustrations, texts, and other components can be associated with one-another automatically. Through manipulation of the structure of the taxonomic elements and through specification of the desired interrelationships, hypotheses about the work can be posed and examined through coordinated inspection of text, illustration, commentary, and bibliography.

Specifically for the Cervantes Project, an XML schema, or modification of the TEI DTD, will be created representing the complex and highly significant interrelationships of episodes and adventures traceable throughout the entire text of Don Quixote as identified and tagged in our narrative taxonomy. Initially only one base text of the Quixote will be fully encoded in TEI XML in the first phase of the textual iconography project, although given the bilingual nature of our site and the international scope of its users, we plan to extend and replicate it also on J. Ormsby’s English translation, already available in our digital library of electronic texts. In the years to come we will investigate scanning and encoding additional key critical/textual editions of the Quixote in XML using the same TEI DTD. Since, this mark-up will include elements created by project staff as modifications of the TEI DTD to represent the various episodes, adventures, themes and motifs present in the narrative, these additionally encoded texts will provide an even richer searching opportunity for Cervantes scholars.
A searchable MySQL database containing metadata for the digital images of illustrations from editions of the *Quixote* in our collection will form the main part of the textual iconography component of the *Cervantes Project*. This metadata will include technical information produced when the images are created, some of which will remain in the archive of master copies with the TIFF as embedded data, to be replicated in the database. Extensive descriptive, structural, and additional administrative metadata for the digitized images of the illustrations, based on the Dublin Core Metadata Element Set will be entered into workforms by project staff. The metadata entered into the image database will also, and most importantly for Cervantes scholars, include terms referring to the appropriate adventures and episodes relevant to the illustrations.

By means of server-side processing, expressed in the PHP scripting language, (although possibly implemented in a more traditional CGI-based architecture or a Java Servlet-based one), metadata for each group of illustrations from a single edition of the *Quixote* will be extracted from the database, and placed within a set of tags representing a ‘mdwrp’ or metadata wrapper element in a METS (Metadata Encoding and Transmission Standard) document. In this way, by means of hyperlinks between different parts of a METS document created for the *Cervantes Project*, according to a structural map developed by project staff, the relationships between all images comprising a single illustrated edition of the *Quixote*, as well as the EAD encoded finding aid and the TEI encoded base text, will be clearly elucidated and displayed. At present, tools for authoring METS documents are in development, as exemplified by Rich Beaubien’s tool for the University of California at Berkeley Library Systems Office (http://www.loc.gov/standards/mets/presentations.html), but the standard is already supported by both the Library of Congress and RLG (Research Libraries Group).

When programmed using METS, all files composing a single digital object in the textual iconography, including TIFF and JPEG image files, EAD XML encoded finding aids, and TEI XML encoded texts, will be bundled via structural metadata allowing for discovery of relationships among the component files. For the *Quixote* textual iconography component of the *Cervantes Project*, an uncompressed TIFF, representing an archival quality master copy image of each illustration accompanying a text, will be produced at standards equivalent to 600 dpi resolution, in 24-bit depth color. These TIFF images will be stored as preservation copies for each image on servers at the Texas A&M University Digital Library, and also will be transferred for manipulation and addition of descriptive and administrative metadata to servers both at TAMU Libraries and the Texas A&M Center for the Study of Digital Libraries.

Derivative image files representing the illustrations, for display and access purposes, will be created by project staff in the JPEG file format using Adobe Photoshop 6.0 software and other custom applications based on open-source software. Some administrative metadata will be produced automatically when the file is opened by project staff in Adobe Photoshop 6.0. Appropriate descriptive and administration metadata elements will also be associated with the image files, possibly entered into work forms created for the project as a part of the programming to implement METS (Metadata Encoding Transmission System), for digital content management and storage. Before
files can be entered into a METS document for the textual iconography project, a system of unique identifiers will be devised for each file that is accompanied by a metadata wrapper element. Note that while JPEG files are convenient for encoding steps and for Web viewing because of their smaller file size and reduced resolution more closely matching screen resolutions, the original TIFF images are also being maintained in stable archival storage. It is likely that we will want to generate new display images in the future from the archival TIFFs as display and network speeds improve. The unique identifiers will allow these new images to be associated seamlessly with the earlier encoded metadata.

The digitized illustrations will be presented for viewing as JPEG file format images accompanied by display of the metadata contained in the metadata wrapper elements. Many of these metadata elements contained in the wrapper for the illustrations, such as subjects and descriptions, particularly of the novel’s recurring themes of character, place, time period, adventure and episode, will correspond to the mark-up of the base text in TEI. By means of the visualization software created for the Cervantes Project by programmers at the CSDL, searching of descriptive metadata elements, keyed to the base text of the Quixote, will be coordinated with searching of descriptive metadata elements included in the METS metadata wrappers provided for all the editions presented as JPEG files of the illustrations. Types of metadata created for the Quixote textual iconography images will include descriptive, structural and administrative elements. These basic metadata elements will be expressed using the Dublin Core Metadata Element Set, including element refinements needed for describing rare materials such as the editions of the Quixote.

METS, a Digital Library Federation initiative, will structure the metadata for the project, so that the proper viewing order and coherence among all the component parts of a single digitized edition of the Quixote, added to the textual iconography of the Cervantes Project, is maintained for web viewing and navigation, as well as for file storage and management. A METS document will provide a structured framework for all the TIFF’s and JPEG’s, as well as any TEI encoded text with its TEI header, and EAD encoded finding guide. METS will be the overarching structural metadata for the textual iconography project. For the Quixote textual iconography component of the Cervantes Project, METS will organize and make possible presentation of each successive page of a single TEI encoded text, along with the accompanying images of illustrations occurring at the appropriate points in the digital display to replicate that of the original paper work. METS, which has been designed to be used with both internal metadata embedded in metadata wrapper (mdwpr) elements, as well as point to external metadata created using other schemas, such as EAD, TEI, and Metadata Object Description Schema (MODS), is also designed to support long term preservation and access to digital objects.

Preliminary description of and intellectual access to the Cervantes Collection at the Cushing Memorial Library and Archives for illustrated editions of the Quixote included in the textual iconography’s corpus will be provided by the creation of an archival finding guide. This finding guide will be encoded in XML, using the EAD DTD (Encoded Archival Description Document Type Definition) encoding schema, version
2002. An electronic finding guide encoded in EAD XML, describing the complete collection of texts and images included in the *Cervantes Project* as a whole might be another option, with the various collections, including that held in the Cushing Memorial Library and Archives, treated more or less as a distinct series. An EAD XML finding guide for the Cervantes Humanities Informatics Collection at the Cushing Memorial Library and Archives will, as a normal part of the work of the archivist in Cushing, be written, encoded and mounted on the TARO (Texas Archival Resources Online) cooperative project web site, so that the series level work for that collection will already be mounted on the TARO servers at the University of Texas at Austin.

For the purposes of the *Quixote* textual iconography, structural metadata will include marking up the full text of at least one copy of the *Quixote* in XML using the TEI (Text Encoding Initiative) encoding schema, specifically version P4, 2002 which can be expressed in XML. The text to be encoded as a base text for the textual iconography, for the purpose of placing, linking and displaying the digital images part of the textual iconography collection, is that now contained in the critical edition of the *Quixote*, edited by Francisco Sevilla Arroyo and Antonio Rey Hazas. This edition was published by the Centro de Estudios Cervantinos in 1993, and is currently included in the *Cervantes Project* web site’s *Cervantes Digital Library* collection of electronic texts: [http://www.csdl.tamu.edu/cervantes/english/ctxt/cec/donq1605.html](http://www.csdl.tamu.edu/cervantes/english/ctxt/cec/donq1605.html) and [http://www.csdl.tamu.edu/cervantes/english/ctxt/cec/donq1615.html](http://www.csdl.tamu.edu/cervantes/english/ctxt/cec/donq1615.html). This HTML encoded digital version based on the first edition of the *Quixote*, published in paper format as part of the *Complete Works* of Cervantes, excludes the originally published notes and bibliography. The overall format of the printed edition of the *Quixote* was maintained in producing the e-text, but the text was processed and divided uniformly into fifty-line pages in order to be able to support the application of the *Cervantes Project* search engine.

This machine-readable representation of a base text of the *Quixote* will be fully encoded in XML using TEI encoding language as a reference for searching the narrative elements represented in our taxonomy which are those interpreted by various artists in illustrations produced for successive editions of the *Quixote* published over the past 400 hundred years.

Encoding this base text of the *Quixote* will require, at minimum, using TEI tag sets for both prose and poetry. Since the TEI Guidelines for XML do not permit concurrent encoding of a text in multiple hierarchies, empty milestone elements and link elements will be used to encode a sonnet, for example, which is printed within a section of prose. Additionally, and possibly most importantly, in order to present the full richness of Cervantes’ scholarship for the textual iconography of the *Quixote*, a set of elements for encoding sections of the text representing the various episodes and adventures generally accepted by Cervantes’ experts and unique to scholarship of the novel will be established [See Appendix VIII for preliminary taxonomy]. These unique elements will be established for the textual iconography project either by means of using existent TEI elements and attribute options, as provided in the Guidelines for Electronic Text Encoding and Interchange, XML-compatible edition ([http://www.tei-c.org/P4X/](http://www.tei-c.org/P4X/)), or by creating a new XML schema to accommodate these unique elements.
A TEI header in the encoded base text will contain any necessary descriptive and administrative metadata. Specifically, the TEI header may include all three types of metadata, including descriptive, structural and administrative. This metadata may be divided into four basic sections: a bibliographic description of the machine-readable text, a description of the way the text has been encoded, a non-bibliographic description of the text or profile of the text, and a revision history. Thus, by transforming the HTML base e-text now on the Cervantes Project web site to TEI XML, all metadata concerning the creation of the e-text, including its revision history will be easily recorded and made available in one document.

As well as creating a project-level MARC record for the textual iconography as an electronic resource, each illustration will also have a cataloging record generated for the TAMU Libraries Voyager OPAC. If not generated through use of MODS, as previously referenced, these MARC records could be generated by means of using Perl, or another programming language, so that a data file based on the metadata created in Dublin Core mapped to MARC21 would automatically be produced and transferred to a file for uploading into the Voyager OPAC.

The added value of the illustrations in the Quixote textual iconography digital archive derive in particular from their innovative treatment and relationship with the collection of electronic text available already at the CP and in the linkages allowing connectivity between images, metadata, and bibliography entries in converted MARC. In addition, the archive provides interactivity between digital images and electronic texts, from different entry points, i.e., browse single images, images with same metadata in a given field (content), sequential images from same edition, or all images related to a particular chapter or adventure, by same artist or in the same edition. We also plan to develop a tool to compare, juxtapose and collage related images from several editions, artists, etc., as part of our research to create new approaches and techniques to display images for analysis, beyond browsing and searching; in that connection we will expand archival description methods, and advance ways in which to integrate texts and images with metadata, as previously done for the images included in the electronic variorum edition of the Quixote using the Interactive Timeline viewer tool (ItLv).

Given the access limitations and inadequacies of previous resources in print, the proposed digital archive based on the illustrations of the Quixote represents a revolutionary leap forward in the preservation and access of rare book collections, while the development and application of new searching aids and interactive visualization tools capable of multiple displays and new analytical approaches significantly increases the scholarly value and educational use of the archive.

Together with the EVE-DQ, the textual iconography digital archive gives users not only broader and better access to more information in a more flexible format, but also places users in control, able to select, connect, display and analyze the visual data at their command for their own individual needs and purposes, as characteristic, of course, of the
Cervantes is one of the world’s greatest authors, widely read and studied. His masterpiece, the Quixote, is a seminal work of great impact and significance and both the book and the character have had tremendous influence throughout the centuries wide influence not just in the creation of the novel as genre and in narrative fiction in general, but in practically all areas of the humanities and fields of study, from art to psychology. It is hard to overestimate the value of providing for the first time access to a digital archive of images containing, documenting and visualizing in a contextual, networked and interactive manner the thousands of illustrations included in the Quixote since its publication in 1605. Students and scholars, curious readers and engaged researchers alike will find in the archive delight and profit, whether their interests are in art, book history, book illustration, literary history, critical and textual studies, narrative theory, cultural studies, hypertext or digital libraries. In addition, the digital archive of images will complement and increase the usefulness of the current elements of the Cervantes Project—textual, documentary and bibliographical—and will in turn add significant value as an educational and cultural electronic resource to the study of Cervantes’ work, culture and time. The result will be a hypertextual digital archive that will provide different types of users, for the first time:

1) access to primary textual resources until now rare or unknown;
2) access to a comprehensive collection of illustrations of artistic, historical, and critical significance, and
3) access to innovative tools to visualize and study images in texts.

In all, the archive is designed to elicit active responsiveness and to provoke new types of inquiries and new forms of textual, visual, and critical analyses. The dissemination through the web of a digital archive about the textual iconography of the Quixote, both at Texas A&M University and through our collaboration with the National Library and the University of Castilla-La Mancha in Spain, will make possible a more complete and profound knowledge of the role, functions and diverse uses of textual illustrations, and will help us understand in particular their contribution as visual narratives and how they have shaped reader’s responses and critical interpretations of the Quixote as an experimental and postmodern text. The proposed digital archive of the textual iconography associated with Cervantes’ Don Quixote is, to our knowledge, in its scope, goals and methods, the first of its kind, and as such we anticipate it will serve as a model for similar iconographic projects.

The digital images, electronic databases, hypertextual archive, and visualizations tools to be created will be fully accessible free of charge through customized interfaces at the portal of the Cervantes Project in the Internet, as well as, in Spanish, through the portals of the National Library of Spain and the University of Castilla-La Mancha. The wide interdisciplinary interest in the Quixote throughout the centuries, its canonical and seminal status in the creation of the novel as a genre, its traditional inclusion in world literature courses, and its iconic status in Hispanic culture, are all factors insuring that the
potential audience for the reference materials to be made available by our proposed project will be large, constant, and varied. It will include scholars in literary and book history interested in evaluating the reception and development of the text—written and visual—, students of the novel and of illustrations researching the role and function of iconography in narrative, and curious readers interested in seeing and appreciating for the first time a rich artistic tradition.
Notes

1 Iconografía de Don Quijote; reproducción heligráfica y foto-tipográfica de 101 láminas elegidas entre las 60 ediciones, diversamente ilustradas, que se han publicado durante 257 años...destinadas a la primera edición de Don Quijote (Barcelona: P. Riera, 1879). Catálogo de la exposición celebrada en la Biblioteca Nacional en el tercer centenario de la publicación del Quijote (Madrid, 1905); Exposición cervantina en la Biblioteca Nacional para conmemorar el CCCXXX aniversario de la muerte de Miguel de Cervantes Saavedra (Madrid, 1946); Juan Givanel Mas, Catálogo de la exposición de iconografía cervantina (Barcelona, 1944). En esta categoría hay que añadir ahora el magnífico catálogo ilustrado de la exposición inaugurada el 30 de setiembre del 2003 en el Museo del Prado basada principalmente en los fondos de la Hispanic Society of America; Patrick Lenaghan, en colaboración con Javier Blas y José Manuel Matilla, Imágenes del Quijote: Modelos de representación en las ediciones de los siglos XVII a XIX (Madrid: Hispanic Society of America-Museo Nacional del Prado-Real Academia de Bellas Artes de San Fernando, Calcografía Nacional, 2003).

2 H. S. Ashbee, An Iconography of Don Quixote, 1605-1895, (Printed for the Author at the University Press, Aberdeen; Issued by the Bibliographical Society, 1895), includes 23 plates; Manuel Henrich, Iconografía de las ediciones del Quijote de Miguel de Cervantes: Reproducción en facsímile de las portadas de 611 ediciones con notas bibliográficas.. (Barcelona: Henrich y Cfa., 1905), 3 vols,

3 I am referring to the 2 vol. edition illustrated by Tony Johannot (Paris: Dubochet, 1836-1837) with 800 xylography vignettes and engravings.


5 The Cervantes Project (CP) is an ongoing long-term project and research initiative dedicated to the development of a comprehensive digital archive based on the works of Miguel de Cervantes. In partnership with the Center for the Study of Digital Libraries and the Cushing Memorial Library and Archives, a division of the Texas A&M University Libraries, our goal is to create an online repository of textual, documentary, bibliographic, and visual electronic resources to serve the needs of students and scholars interested in Cervantes’ life, times and work, and focused in particular on the study of Don Quijote de la Mancha (http://www.csdl.tamu.edu/cervantes).


8 The Cervantes Chair was established in 2003 with funding from the Banco de Santander (Spain).


10 It has recently come to our attention that in collaboration with the Centro de Estudios Cervantinos (Alcalá de Henares) the Centro de Estudios Cervantinos en Guanajuato (México) is planning to develop a “Banco de imágenes del Quijote” initially based on the 800 paintings, drawings and sculptures at the Museo Iconográfico del Quijote, and to include in the future illustrations from the library donated to the Center by Eulalio Ferrer.