The Electronic variorum edition of ‘Don Quixote’ at the Cervantes Project

Eduardo Urbina, Richard Furuta, Carlos Monroy, Neal Audenuert, Jie Deng, and Erika Pasquel
Texas A&M University

The Cervantes Project hosted at Texas A&M University’s Center for the Study of Digital Libraries (http://cervantes.tamu.edu), began in 1995 with the creation of an Electronic Variorum Edition of Don Quixote (EVE-DQ)—the first one of its kind to be published electronically—as one of its goals, in the context of re-evaluating traditional scholarly practices, developing new forms of access to primary materials and creating alternatives to the critical apparatus.

The EVE-DQ is a collaborative project started in 1998 by a team of researchers in Humanities and Computer Science at Texas A&M University’s Center for the Study of Digital Libraries. It received initial research support from the Office of the Vice President of Research and the College of Liberal Arts, was funded in 2000 by a three-year grant from the National Science Foundation programs in Digital Libraries and Information Technology Research, and since 2003 receives support from the Cátedra Cervantes of the Universidad de Castilla-La Mancha and the Digital Humanities program at Texas A&M University.

The focus of the EVE-DQ is the application of advanced computer science research to update and enhance traditional scholarly practices. Our primary goal is to develop a replicable program that permits the creation of online critical editions as hypertextual archives, using the Quixote as testbed. Our work addresses the future of the tradition of bibliographical and textual studies established by Greg, Bowers, and Tanselle, and proposes the creation of virtual electronic editions by direct access to online hypertextual archives and databases.
The number of theoretical and programmatic studies envisioning new types of texts and editions, or proposing new paradigms of electronic textuality far exceeds the actual examples of developed projects and scholarly collections of critically edited literary electronic texts, freely accessible to all users. The creation third-generation digital collections of literary texts such as the works of Chaucer, Shakespeare and Cervantes remain complex, expensive, and difficult enterprises caught between tradition and innovation. They require for their successful production and existence a delicate combination of original collaborative scholarship, groundbreaking interdisciplinary research, and forward looking financial and institutional support, such as one can find and appreciate in The Canterbury Tales Project, the Shakespeare Electronic Archive and the Perseus Digital Library, to cite only some of the most notable collections and projects.\(^5\)

Textual scholars and editors have been slow to react to the new electronic paradigm, and even more so in recognizing the sure impact and profound implications of the Internet on their discipline. Thus, even such pioneering minds and hypertext practitioners as Landow and Robinson were still addressing in 1996 the technical problems posed by annotating hypertexts and the future direction of hyperediting.\(^6\) And it was not until the 3\(^{\text{rd}}\) edition of Shillingsburg’s Scholarly Editing in the Computing Age (1996), for instance, that a chapter on electronic editing was included, and even then it was infused by a serious doses of skepticism as to the merits of non-linear editions.\(^7\)

The existence of a mere dozen or so major electronic publishing projects, according to Robinson, seems to reveal precisely that “the power of the computer to search, to sort, to navigate, to explore” remains an unrealized promise, as his own proposal for a future model and methodology would confirm.\(^8\) Furthermore, in 1992 Richard Knowles stated, regarding the then contemplated possibility of the MLA publishing the new variorum Shakespeare both in book and
Siemens, Cook, Hockey and others have dared to anticipate and predict the future of hypertextual archives and to envision, even, the coming of a new type of electronic editions capable of encompassing and providing access to all available textual knowledge and scholarship—graphical, critical, and historical. Our project accomplish such a goal by developing the end-to-end program and tools capable of converting, collating, annotating, and composing, within the classical tradition of textual studies, multiple scholarly editions, from a documentary edition to a critical *variorum* edition.

According to Paul Wernstine, “there has never been an old-spelling edition of Shakespeare’s so called canon prepared according to the principles of copy-text editing . . . definitive and entirely objective.” Such impossible quest seems to have dissipated for now in the age of poststructuralist theory and textual deconstruction, as a fatal doubt was cast on the idea of reconstructing authorial meanings out of non-existent texts. In fact, the rise of multiple interpretative communities contributed in part to the development of new editorial practices in the 1990s which include, as a matter of fact, hypertext editions. Notably, as Wernstine observes again speaking of Shakespeare’s works,

Hypertext allows the Shakespearean editor finally to confront in practice the knowledge that since we do not know anything specific about the manuscript sources of the printed plays, we never had any grounds (besides aesthetic or practical ones) for choice between one printing of a play or another. . . . The challenge for editors shifts from the task of generating editorial theory that can be deployed in the defense of archaic editorial choices to the practical task of how to display in book form for a wide audience the indeterminacy that characterizes Shakespeare textuality. (52)

The *EVE-DQ* is not, strictly speaking, an electronic edition but rather a dynamic hypertextual archive composed of a series of databases—graphical, textual, and documentary—and two principal modules, the editing module or MVED (Multi-variant editor of documents),
and the visualization-composition module or VERI (Virtual edition readers interface), which in turn are controlled by a hypermedia-based data entity management system (HDEMS). The initial process of collation involves the creation of data entities. For example, a variant that is being annotated is a data entity. The HDEMS manages these data entities by preserving the relations and linkages among them, and maintaining consistencies amongst related data entities. The HDEMS application interface (API) keeps track and manages within the MVED the relational databases created in the collation/editing/annotation/emendation process. All other modules call the HDEMS API functions and the actual interaction with the databases is done by the HDEMS API.

The editing program or MVED, is comprised of seven tools or modules, which can be briefly described as follows:

1. The **Collator program** automatically identifies textual variants among two or more of the selected editions/copies against a base text.
2. The **Index of Variants** generates multiple lists of textual variants identified during the collation and provides interactive visual representations of such variants.
3. The **Text-Image Synchronizer** uses predetermined synchronization points between the digital images and the electronic texts for each of the editions/copies included in the collation process.
4. The **Dual-form Document Viewer** shows together the synchronized texts and associated images to be able to browse the source editions incorporated into the collation.
5. The **Editing/Emendation tool** allows the editor to correct and annotate both variants and non-variant text, and creates the hyperlinked data entities stored in the relational database.
6. The **Data Entity Browsing Interface** provides the editor with updated access to all the details and information produced during the editing process.
7. The **Annotation Tool** allows the classification and annotation of variants as well as the inclusion of critical commentary and references during the editing process of any of the collated texts.
Essentially the MVED is a collation tool. Modules such as the “Collator program,” the “Lists of variants,” the “Text-image synchronizer,” and the “Dual-form document viewer” function within the collation process but are independent of the editor being logged in, while the “Data Entity Browsing Interface” is used only within the collation process. One can, however, browse through documents using the browsing interface. Thus, although all the tools are used in the collation process, the “Dual-form document viewer” and the “Text-image synchronizer” can be used independently. In addition, the “Editing/Emendation tool” and the “Annotation tool” are designed to be used together so that one can annotate while editing a variant, or while doing an emendation. This dual function is only possible during the collation process, and it is necessary that a valid editor be logged in to the system.

The readers or user’s module is called VERI or Virtual Edition Readers Interface. This module connects with the EVE’s textual, graphical and relational databases stored in the MVED, provides browsing and searching functions and allows users to compose individualized editions from the collections of digital texts, collations, and interlinked hypertextual annotations. The VERI is also an interactive tool capable of providing access to individual editions, textual or graphical and allowing selection of editors, levels and types of annotations in order to create what we call a virtual critical edition. Specifically, this tool makes possible:

1) Selective access to and visualization in different modes of electronic and digital texts and hypertextual links produced by the collations performed by the MVED
2) Synchronization of texts and images and visualization in different resolutions of images for closer analysis of texts and variants
3) Full text search engine of all electronic texts included in the *variorum*
4) Individualized composition of editions based on reader’s preference and needs using the base texts, variants, and textual documentation and annotations resulting
from the operation of the seven modules of the MVED and included in the EVE; from a single digital facsimile to a critical edition

During the last seven years, in the process of designing, developing, and deploying programs and tools, our extensive discussions have convinced us of the fundamental changes to the traditional elements of critical editing and textual analysis that the new forms and types of access to the texts make possible. The first step in the creation of the *EVE-DQ* was the development of the MVED (Multi Variant Editor for Documents). The MVED enables the collation of multiple texts. The variants are stored in a relational database. Editors can then classify the variants, write annotations, and introduce critical commentary to the texts, having instant access to both the textual transcription and the facsimile of the original copy. In essence, the MVED is a fundamental change in the way scholars collate and annotate a literary work. Since there is no “final version,” editors can dynamically modify editing decisions and add new commentaries and annotations. Corrections are reflected immediately in the electronic edition, for instance, whenever editors discover—after further analysis on the source texts—that a variant has been wrongly classified. In creating the *EVE-DQ*, we found that facsimile copies of the original texts continued to become available as agreements with their owners or holders were reached. We would not have been able to include these in a printed publication, but have been able to include them incrementally in the *EVE-DQ*.

Additionally, the *EVE-DQ*’s dynamic hypertextual environment introduces extensive and fundamental changes to the reader’s access, participation and role, subsuming to some degree the functions of editor and publisher. In a printed edition, readers cannot look into the source texts; they can only access the results of the collation, which is the only way for them to browse and read the critical *variorum* edition. This version, although good enough in providing an overview
of the variants, does not allow readers to have different views and perspectives of the result of the collation, limiting the ways the edition can be browsed.\footnote{13}

Readers access the \textit{EVE-DQ} through the VERI (Virtual Edition Reader’s Interface), which enables readers to access the source texts and to compose their own editions. For instance, readers can filter out some variants based on certain categories, thus narrowing the browsing scope of their information space. Readers’ browsing of the \textit{EVE-DQ} also differs from a printed edition. The \textit{EVE-DQ} presents readers with the base text as the starting point. Color highlights identify the segments in the base text that correspond to a variant in any of the other comparison texts. Readers also can display the variants either alphabetically or by the order in which they appear in the text.

From the reader’s perspective, reading a printed variorum edition is driven by the way in which the edition was composed by the editor and publisher. Clearly, this approach limits the reader’s options. Volume III of \textit{Don Quixote’s} edition published by John Bowle in London in 1781, for example, includes the results of collating the Madrid 1605 and 1608, Valencia 1605, and London 1738 editions.\footnote{14} Variants in this volume are presented in columns, where each column corresponds to one of the source editions, and are listed in the order in which they appear in the text, grouped by chapter. Annotations are presented in a separate volume, with only a reference associated with the corresponding text. While retaining much of the physical form of the printed volume, electronic presentation enables reshaping the page order to bring related elements into physical proximity.

The architecture of the EVE-DQ also enables further experimentation on ways to present collation results. We have started preliminary tests using an interactive timeline-based visualization tool, ItLv (Interactive Timeline Viewer).\footnote{15} This tool provides readers with
additional ways to visualize the results of the collation, and to explore and analyze them in order
to find patterns and correlations among the texts.

Our work on the EVE-DQ has taken us to the realization of creating not an electronic
edition but a series of interlinked digital archives forming a hyperedition. This approach
coincides with the one presented by Donaldson in the *Shakespeare Electronic Archive*.16 He
expands the notion of a text, showing how a hypertextual edition captures, represents, and makes
accessible the rich graphic and textual tradition of Shakespeare’s works, calling it a “living
variorum . . . an electronic archive.” Donaldson describes the main differences regarding the use
of the texts, their presentation and access, and the new nature and functionality of a hypertextual
edition,

In contrast to the modern critical edition in print, in which the main text is a
distillation of the evidence, and only one of many that are possible, the
documentary hypertext presents images of the evidence itself, and does so in a
dynamic structure that makes it possible to toggle rapidly between alternate states
so that not merely the fact and the content of a variation can be noted, but the
effect of alternatives can be experienced in the context of the text in which it
appears. (186)

The characteristics implemented in the development of the *EVE-DQ* as documentary
hypertext or hyperedition lead us to reconsider our initial premises regarding the focus on the
editor and his authorial control of the critical apparatus and to implement instead a text-central
approach based on direct, immediate and dynamic access to digital facsimiles, electronic textual
transcriptions, and collated variants through hyperlinks, image-text synchronization and
relational databases. In turn, we have had to reconsider the three key elements of the critical
edition, *recensio, collatio* and *emendatio*, and how they are affected by the new forms of access
and presentation in hyperediting.
Williams and Abbot define *recensio* as the reconstruction of the lost common ancestor of surviving texts through the examination of manuscripts and the texts they contain. The digital environment of the EVE-DQ allows not just the examination and reconstruction of texts in a selective and exclusive fashion by a privileged and authoritative editor, but the inclusion and visualization in an electronic hypertextual archive of all the surviving copies and editions, both in image and textual format, for the examination of any editor or authoring user. This level of access to textual information changes the rationale and function of the critical apparatus in two ways. First, complete, universal, and immediate textual access makes superfluous to a large extent notes and other items describing the accidentals of the texts, since the editor is no longer driven by the need to reconstruct and publish a unique or single edited text. Second, readers can visualize the text of any particular copy or edition without the need to accept the hidden choices or personal conjectures of the editor. The relationships between the text (author) and editor, the role of the editor as agent of knowledge have changed. Similarly, the role of the reader, is no longer a passive receptor of filtered or absent documentation, but instead a fully informed and active participant.

In textual criticism, *collatio* is the operation by which one text is compared to another text to discover textual variation. This time-consuming and tedious process has benefited most from technology and computer applications. In our case, the MVED allows the collation of two or more texts and offers a complete flexible and interactive index of variants, the ability to automatically create links from the variants to the texts, and to synchronize them with the digitized images of the textual sources. In addition, it offers a mechanism to edit and annotate variants and non-variant text, and to store all annotations, commentaries and references in a
Finally, the application of new methods and tools has also had an impact on the \emendatio, the use of conjecture to emend the text reconstructed from surviving witnesses, thereby removing errors. With the use of the MVED the need for \emendatio if not altogether eliminated, becomes greatly diminished. Since the number and types of conjectures introduced by the editor are often driven by a limited collation process and predicated by the traditional methods of the \recensio, it is thus clear that the paradigmatic effects caused in such critical processes by the new technology and electronic tools have and will continue to have from now on profound consequences in the creation of hypereditions such as the \textit{Quixote variorum}.

In the nineties, well known literary works began to experience new presence in the digital world. This presence was driven—among other things—by the development and advancement of new information technology tools and techniques, especially the World-Wide Web. In this context, Cervantes’ works in general and \textit{Don Quixote} in particular were not ignored. In 1996, Fred Jehle published the first electronic edition of the \textit{Quixote} to appear on the Web based on Schevill & Bonilla´s critical edition.\textsuperscript{18}

Since 1998 our work has been focused on the development of tools and programs to create pure electronic editions, that is to say, editions that from the start have as a goal the production of electronic texts using digital means and information technology research. In our work, we take advantage of currently available technologies and advances in electronic editing, and follow the new developments in textual criticism proposed since the early 1990’s by Charles Faulhaber, George Landow, and Peter Robinson among others,\textsuperscript{19} conceptualized by Jerome McGann in his trailblazing study “The Rationale of Hypertext,” and exemplified in his project,
These new forms of textuality aim to produce hypertexts, hypereditions, and hypertextual archives of the type anticipated by Landow, envisioned by Faulhaber, Marcos Marín, Hockey, and Lavagnino, and initiated by Robinson, Donaldson, and McGann. As far as the development, production, and publication of complex scholarly editions, such as *variorum* editions, and as exemplified in our edition of the *Quixote (EVE-DQ)*, the future of such texts as archives is most certainly electronic, networked, and hypertextual.

**Notes**


12 The initial design of the VERI and its functions was presented in Spanish, along with a first version of this study at a course on Digital Libraries and University Libraries organized by the University of Castilla-La Mancha in November 2002. The first version of the EVE-DQ was published and presented in June 2005 at the Biblioteca Nacional in Madrid and at the University of Castilla-La Mancha, and it can be accessed at http://www.csdli.tamu.edu/cervantes/V2/variorum/index.htm.


14 John Bowle, Anotaciones a la Historia de Don Quijote de la Mancha (London-Salisbury, 1781); now available at the Cervantes Project Digital Library: http://csdli.cs.tamu.edu:8080/veri/indexRefs.html


ABSTRACT The Cervantes Project is creating an Electronic Variorum Edition (EVE) of Cervantes’ well-known Don Quixote de la Mancha, published beginning in 1605. In this paper, we report on visualizations of features of a text collection that help us validate our text transcriptions and understand the relationships among the different printings of an edition. Categories and Subject Descriptors H.3.7 [Digital Libraries]; J.5 [Arts and Humanities]. General Terms Design, Experimentation, Human Factors. Some interesting points come out when we look at the visualizations. Consider Figure 2. The display shows that the base text and texts 1, 2, 3, and 4 are similar to each other, since the number of variance points is not too large. The base text and copy 5 are different from each other. Size: px. Start display at page: Download "An Electronic Edition of Don Quixote for Humanities Scholars". Error 2 Choosing one variance over another is the job of a Cervantes scholar, and different scholars may, of course, make different choices based on their interpretation of the raw material. Our acquisition process supports the specification of such choices, allowing the categorization of the choice and association of justification for it. The virtual variorum edition will contain multiple copies of the 11 significant early editions of Don Quixote, annotation of the variances present among the editions to allow their comparison, derivative editions, generated as the result of scholarly analysis of the variances and bearing supporting reasoning, and scholarly commentary by experts. Don Quixote, Part II, emerged from the same press as its predecessor late in 1615. It was quickly reprinted in Brussels and Valencia, 1616, and Lisbon, 1617. Parts I and II first appeared in one edition in Barcelona, 1617. There was a French translation of Part II by 1618 and an English one by 1620. The second part capitalizes on the potential of the first, developing and diversifying without sacrificing familiarity.