

Scholarly Digital Editions: A Wise Investment for Scholars and Institutions

Samuel J. Huskey

Abstract This author advocates for the publication of digital scholarly editions. After examining what constitutes a digital scholarly edition, he continues by arguing that the digital paradigm presents an opportunity for separating the content from the presentation or visualization of an edition, and furthermore that textual scholars should not be expected to design digital interfaces. The second section considers how the Digital Latin Library project is implementing this concept in its series “The Library of Digital Latin Texts”, which seeks to provide an outlet for publishing peer-reviewed scholarly digital editions as version-controlled TEI XML databases.

Zusammenfassung Der Verfasser plädiert für die Erstellung digitaler kritischer Editionen. Nach einer Erörterung, was eine digitale kritische Ausgabe ausmacht, ist er der Meinung, dass das digitale Paradigma die Möglichkeit bietet, den Inhalt einer Edition und ihre Präsentation/Visualisierung auseinanderzuhalten. Dies bedeutet, dass von den Editoren nicht erwartet wird, die digitalen Interfaces einer Edition zu entwerfen. Der zweite Teil prüft, wie das Projekt Digital Latin Library dieses Konzept in der Reihe „The Library of Digital Latin Texts“ einsetzt. Ziel der Reihe ist es, eine Publikationsplattform für peer-reviewed digitale kritische Editionen in Form von versionkontrollierten XML-Datenbanken zu sein.

Keywords critical editions, visualisation, apparatus criticus, sustainability, technical knowledge

To answer the question of why scholars and institutions should invest in scholarly digital editions, let us begin by affirming the reasons for investing in scholarly editions at all, digital or not. Scholarly editions present reliable texts.¹ They assemble and preserve historical information related to a text. They aim to provide reliable, trustworthy guidance to the unique issues and problems inherent in a text and

1 This is, in fact, a paraphrase of the first sentence of the Modern Language Association’s “Guidelines for Editors of Scholarly Editions” (2012).

its contexts. Most important of all, they contribute to the scholarly conversation about the text and its tradition. How the wider community of scholars values an edition determines its overall impact—and how soon another scholarly edition of the same text might challenge its position. Even after they have been superseded, scholarly editions can continue to be useful as reference works, sometimes for decades or even centuries.² The march of scholarship ever onward thus makes scholarly editions wise investments, since they can influence the discussion for the long term.

It remains to be seen, however, if scholarly digital editions are prudent long-term investments. It would seem to be a simple case to make: if scholarly editions are worthwhile, then scholarly editions that can leverage massive amounts of data and processing power to provide new insights into texts and their contexts are surely worthwhile, too. But rapid changes in technology, high production costs, the challenges of maintenance and upkeep, lack of consensus regarding best practices and standards, and disagreements about what even constitutes a scholarly digital edition have put the investment into the “high risk” category.³ Nevertheless, risky investments can be highly profitable. If we can achieve some clarity about the nature of scholarly digital editions and our expectations of them, we could reap dividends for the foreseeable future.

Crass financial metaphors aside, the language of investing is not out of place here, considering the significant sums of money, resources, and time that scholarly digital editions often require. Although funding agencies have generously supported efforts to explore the potential of scholarly digital editions, that support will not continue without an obvious return on investment. Similarly, scholars and the institutions that employ them must be mindful stewards of their resources, so giving some thought to the benefits and risks of investing in scholarly digital editions is prudent. Accordingly, this chapter aims to make a case for limiting the risks inherent in scholarly digital editions and maximizing their profit for the scholarly community. It will begin by challenging the current notion of what a scholarly digital edition is by reviewing some of the issues and problems in this sector of scholarship. It will then present as a case study the *Digital Latin Library*'s series the *Library of Digital Latin Texts* as a potential model for a stable, long-term investment in scholarly digital editions.

2 Consider the fact that Owen's edition of Ovid's *Tristia*, famously savaged by Housman (1915), continues to be cited in modern scholarship, despite the existence of two vastly superior editions by Luck (1967) and Hall (1995), with ongoing debate about which of those two is preferable.

3 See Del Turco (2016) for an extended discussion of these issues.

1 What is a Scholarly Digital Edition?

What is meant by “scholarly digital edition” has long been the subject of debate and discussion.⁴ The *Frequently Asked Questions* page of the *Catalogue of Digital Editions* effectively illustrates the problem.⁵ In response to the question “What are the criteria for inclusion in the Catalogue?” the editors explain:

We primarily look for digital editions and digital scholarly editions (by which we mean editions with a strong critical component) but people are also beginning to submit digital archives, textual collections where some texts are treated in more detail, etc. It’s worth noting that people define their projects in various ways (e.g. a project may define itself as a database when in fact it might be more of a digital archive) using different terms synonymously. This makes it difficult to pick what should be included in our list or not. The line of inclusion in our catalogue is a little blurry but it reflects the fuzziness that comes with these projects.

Browsing the catalog itself reveals a wide variety of projects, with over 250 items describing themselves as archives, databases, variorum editions, portals, etc. Fortunately, the catalogue’s search feature offers “Scholarly,” “Digital,” and “Edition,” as filters:⁶

- Scholarly: “An edition must be critical, must have critical components—a pure facsimile is not an edition, a digital library is not an edition.”
- Digital: “A digital edition can not (sic) be converted to a printed edition without substantial loss of content or functionality—vice versa: a retrodigitized printed edition is not a Scholarly Digital Edition (but it may evolve into a Scholarly Digital Edition through new content or functionalities).”
- Edition: “An edition must represent its material (usually as transcribed/edited text)—a catalog, an index, a descriptive database is not an edition.”

That these definitions are negative is telling: upon surveying the cornucopia of projects in the catalog, it does appear to be easier to say what is not a scholarly digital edition than what is. Even after applying those filters, the diversity of projects and their platforms frustrates efforts to define “scholarly digital edition” with much precision.

4 E.g., Robinson (2002), Ducourtieux (2004), Gabler (2010), Sahle (2007), Sahle (2013), and Sahle (2016).

5 Franzini et al. (2016–) <https://dig-ed-cat.acdh.oeaw.ac.at/faq>, accessed 15.01.2018.

6 Franzini et al. (2016–). The categories are based on Sahle’s work on defining scholarly digital editions. See Sahle (2016) for a summary.

Sahle discusses some previous definitions of scholarly digital editions before proposing one of his own: “Scholarly digital editions are scholarly editions that are guided by a digital paradigm in their theory, method and practice.”⁷ There is much to like in Sahle’s discussion, but his definition puts too much emphasis on the digital aspect of scholarly digital editions. That might seem to be an absurd objection. After all, it makes sense to require that a digital edition be fundamentally digital. But why must there be a “digital paradigm” for their theory, method, and practice? The “digital turn” has not rendered irrelevant the traditional theories, methods, and practices of scholarly editing; they are relevant regardless of the medium in which they are applied. Indeed, the problem with Sahle’s definition, and others like it, is that it confuses matters by tying the edition to its presentational medium. That is, it makes editors responsible for both the quality of the text and the design and features of the interface for making it available to readers—a quantum leap from our expectations of editors of traditional editions.

We expect editors of volumes in the venerable *Bibliotheca Teubneriana*, for example, to deliver a high quality text with a preface, critical apparatus, and other components that reflect their discernment and good scholarly judgment, but we do not also expect them to design the typeface and the layout for the main text and the apparatus criticus, make the paper, sew the binding of the book, or conduct marketing and distribution efforts. Yet, current expectations for editors of digital editions extend to issues of design, user experience, information architecture, encoding standards, file formats, accessibility, and a range of other technical matters that lie beyond the realm of philology and the fundamental purpose of editions as articulated in the opening paragraph of this chapter.

To meet the expectations of a digital paradigm, editors of scholarly digital editions would have to master not only the intricacies of textual criticism and the idiosyncrasies of their chosen text, but also the full stack of technology required for designing, building, publishing, and maintaining a digital resource. To be sure, some scholars have managed to become proficient in all of these areas, and others have been successful in obtaining funding to pay people to design, build, and maintain the digital components. Nevertheless, it is not just unreasonable, but unsustainable to ask textual scholars to master the wide array of technology that most industries hire teams of professionals to implement and manage. Planned obsolescence, rapid development of new technologies, and the normal lifecycle of hardware and software make it a given that scholarly digital editions that rely on specific technology will eventually cease to function as expected if they are not maintained and updated.⁸ How many textual scholars have the time

7 Sahle (2016) 28.

8 Del Turco (2016) 228 observes, “While manuscripts may have lasted hundreds of years, it is disconcerting to note how the life span of a digital facsimile/edition is sometimes less

and resources to stay current with the rapid developments of an entirely separate technical field?"

It is worth adding that the same obstacles bar the path to effective peer review of scholarly digital editions. It can be difficult to find reviewers for scholarly editions of any kind, but identifying reviewers who can also competently assess digital components can be nearly impossible. To mitigate the problem, the Institute for Documentology and Scholarly Editing has developed the online journal *RIDE* (*Review Journal of the IDE*), but their "Criteria for Reviewing Scholarly Digital Editions" highlights the nature of the problem.¹⁰ The first three sections of their criteria address standard elements that will be found in a review of any scholarly resource, digital or otherwise: Bibliography, Content, Method, etc. The fourth section, however, introduces technical matters that call for specialized knowledge: Technical Infrastructure, Interface and Usability, Browse, Search, Metadata, Social Integration, Rights and Licenses, etc. Although reviewers of printed publications sometimes remark on the quality of the physical product, it is hard to imagine a reviewer of a traditional edition holding the editor responsible for analogous aspects of the codex platform (i. e., a book). The response to this criticism, of course, is that these technical matters cannot be separated from the arguments advanced by scholarly digital editions. But if that is the case, it is worth asking if scholarly digital editions are actually conglomerations of multiple forms of scholarship that complement each other, but also deserve to be considered as individual contributions to the discussion.

2 Scholarly Digital Edition or Data Visualization?

Many, if not all, of the criteria used to distinguish scholarly digital editions from traditional printed editions have to do with how the digital format conveys information. Reviews in *RIDE*, for example, devote considerable attention to navigation and presentation of information. Gengnagel's review of an edition of Goethe's *Faust* is a good example:

than 4–5 years." See also McGann (2013), 286: "[...] most—nearly all—websites created in HTML will not outlive their creators, and the duration of the materials may well be much shorter even than that."

9 Gailey/Jewell (2012) 5–6 coin the term "hipster ethos" to refer to the opposite phenomenon, in which "the quality of the work you do is not so important as staying at the edge of innovation, always one step ahead of the unfashionable masses."

10 <https://www.i-d-e.de/publikationen/weitereschriften/criteria-version-1->, accessed 27.12.2017.

As a last point in this section on aesthetics and navigation, I would like to point out that the visualization of the variants in the reading text is inspired. Indicating that a line has variants in other witnesses and indicating the quantity of variants by colouring the background of the lines increasingly darker, the more variants there are, is a visualization that is readily apparent in itself. The possibility to toggle the variants directly in the text is a welcome utilization of the hypertextual concept underlying web-based editions as opposed to the footnote apparatuses used in printed editions out of a certain necessity.

The fact that she uses the word “visualization” twice in this excerpt is instructive, since that is, in fact, the crux of the matter. As a scholarly endeavor, data visualization refers to the creation and study of representations of information. Practitioners use artistic, philosophical, psychological, and scientific methods and techniques to understand how we relate to, process, and use information, and they use that knowledge to develop new ways of visualizing data. Given that the digital aspect of many scholarly digital editions is concerned with the display of information, and given that reviews of scholarly digital editions consider how the user experiences the information, it seems worthwhile to ask if scholarly digital editions as currently conceived are, in fact, single products of scholarship.¹¹

We are so accustomed to the conventions of scholarly editions in print that we do not recognize traditional editions for what they are: visually encoded collections of curated information. That is, they rely on symbols, abbreviations, typographical conventions, and page layout to present arguments about complex, structured information. Considered in this way, scholarly editions are databases, but their data are frozen on the printed page and inaccessible to digital tools for display and analysis. Nevertheless, the format of an edition has shown itself to be a highly useful and efficient way to present textual information to readers. But the book is only one platform for storing, retrieving, and using a critical edition. Moreover, the traditional layout and typographical conventions of a scholarly edition are only one way of presenting its information. One could call it a print-optimized visualization of textual data.

Most decisions about the layout and design of printed publications are determined by house style guides and the physical constraints of the page. Editors of printed editions can have some influence over matters of layout and design. The various critical editions of Servius published over the years provide instructive examples of editors stretching the limits of layout and design to communicate details

¹¹ Gabler (2010) 50 anticipates this argument by discussing visualization as one element in the “web of discourses” found in a scholarly digital edition.

about the text in a visual way.¹² Most often, scholars develop reference systems, abbreviations, and symbols to work within the constraints of their publisher's guidelines.¹³ In any case, how a printed edition looks is up to people other than the editor (e.g., typesetters and compositors), and their job is to limit their work's impact on the edition's scholarship. Since they are not making any innovations in the manner of presentation, but rather applying a style guide to make a product, their work is not scholarship, but a craft. But for some reason that has not been the model for scholarly digital editions. We continue to hold editors of scholarly digital editions accountable for the content of the edition, but we also expect them to design and implement an information architecture, paying special attention to the visualization and display of the edition's data. That has the effect of devaluing the scholarly work of both editing and data visualization, since it suggests that an editor should be able to pick up the skills required for information management and visualization, or that a scholar of data visualization should be able to pick up the skills required for scholarly editing. The fact is that editing and data visualization are distinct fields of scholarship, and they should be treated as such. Simply put, the editor's task should be to deliver the highest quality text and ancillary materials as possible. How that information is delivered to the world is an entirely separate, but vitally important area for scholarly endeavor.

This is not to suggest that it is impossible to do both scholarly editing and data visualization, or that a scholarly digital edition cannot be the product of a fruitful collaboration between scholars in those fields. Rather, it is an appeal to reconsider our expectations of scholarly digital editions and to value the philological, visual, and technical components as scholarly endeavors in and of themselves.

The next section demonstrates one way of doing that. The point of reference will be the *Library of Digital Latin Texts (LDLT)*, the collection of digital critical editions that the *Digital Latin Library (DLL)* aims to publish in collaboration with a number of learned societies.

3 The *Library of Digital Latin Texts*

The *DLL* project began in 2012 with a two-fold mission: 1) Publishing and curating critical editions of Latin texts, of all types, from all eras, and facilitating an ongoing scholarly conversation about these texts through open collaboration and

12 See Murgia (2018) xx–xxviii for a discussion of the various methods that have been used to represent Servius' commentary on the printed page.

13 McGann (2013) 282–283 discusses the intricate reference system in Mays (2001) to show that linked data can occur in printed editions, too.

annotation; 2) Facilitating the finding and, where openly available and accessible online, the reading of all texts written in Latin. The *LDLT* serves the first part of the mission through a collaboration with three learned organizations: The Society for Classical Studies (SCS), The Medieval Academy of America (MAA), and the Renaissance Society of America (RSA). Through their organizational structures for publishing scholarship of interest to the communities they represent, these learned societies are currently developing or finalizing policies and procedures for publishing editions in the *LDLT*.

4 Separating Content and Visualization

A defining feature of the *LDLT* is that it is a library in the sense of a uniform collection of volumes. The reasons for this emphasis on uniformity were announced in a blog post explaining the nature of the *LDLT*:¹⁴

Trying to accommodate every vision for a multimedia edition would rapidly exhaust the *DLL*'s resources and practically guarantee that the *LDLT* would be a repository of unique, isolated projects. Instead, we aim to publish a uniform collection of texts that can be stored, retrieved, viewed, queried, and analyzed with reliable, predictable results.

This is not meant to disparage projects that leverage the availability of multiple media and different kinds of technology to highlight unique characteristics of a text or a collection of texts. Such projects can be useful and important contributions to scholarship. Rather, the *LDLT* aims to provide a platform for publishing structured textual data in an open and accessible way that will support other efforts to promote the understanding of texts and their contexts.

To highlight the importance of visualization and the creative reuse of structured textual data as forms of scholarship in their own right, a major portion of the *DLL* project has been to study how people interact with traditional and digital versions of texts, and to develop candidate techniques for the display and visual analysis of *LDLT* texts in particular. June Abbas, a co-PI on the *DLL* project and a scholar of library and information science, analyzed how users interact with critical editions and produced a user behavior study that has guided much of the *DLL*'s work in this regard.¹⁵ Chris Weaver, another co-PI and a scholar of data visualization,

¹⁴ Huskey (2017).

¹⁵ Abbas et al. (2015) and Abbas et al. (2016).

has drawn on Abbas' work to develop a number of visualizations of critical edition data, ranging from basic displays of text to dynamic, graphical representations of the text and its metadata. A desktop application for working with these visualization techniques will be available for downloading from the *DLL*'s site after the publication of the first *LDLT* editions.

So that there will be a human-friendly way of interacting with the data of an *LDLT* edition, Hugh Cayless, of the Duke Collaboratory for Classics Computing, has developed the *DLL*'s official edition viewer, a web-based reading environment that enables a number of dynamic features, such as live swapping of variant readings into and out of the text, filtering of different types of variants, clickable links for the display of some source materials, and tools for lexical and morphological analysis.¹⁶ Although some might consider this application to be "just" a tool, Cayless' work is in fact a good example of why we need to treat scholarly editing and interface development as separate, equally important works of scholarship. Cayless' edition viewer is a work of scholarship in that it makes an argument about how the discrete pieces of information in an edition relate to each other and why that is important. The fact that it also facilitates the reading of *LDLT* editions is a benefit of Cayless' scholarship, and its adoption by the *DLL* as its official reading application is an affirmation of its contribution to the field.¹⁷

Without continual development and updating, these tools for reading and visualizing *LDLT* editions will inevitably break. That is the nature of technology—it is also another argument for separating content from display. For that reason, an *LDLT* edition is simply a single text file encoded in XML according to the *LDLT*'s guidelines and stored in a version-controlled repository. It may contain links to external resources and to resources stored alongside it in the repository such as collation tables, transcriptions, notes, images, etc., but an *LDLT* edition is published as a single file that contains the text and information about it (e. g., preface, text, critical apparatus, etc.). Because the data is openly available, the editor or anyone else is free to use it as the basis for separate visualization projects to be published elsewhere. And because the file is in essence a text file, its data can be migrated to a different platform should XML become obsolete at some point in the future. This allows the editor to focus on editing, without having to think about how to display the text or trying to keep up with the technical issues of file formats or information architecture.

16 Code available at <https://github.com/DigitalLatin/viewer>, accessed 23.01.2018.

17 See Cayless (2018) for more discussion of interfaces as interpretations of data models.

5 Peer Review and Publication

Aside from this issue of recognizing editing and visualization as separate, but complementary, forms of scholarship, the *LDLT* also aims to address the paucity of viable outlets for peer-reviewed publication of scholarly digital editions. Journals such as *RIDE* already exist to provide a forum for reviewing existing scholarly digital editions. These independent reviews serve the vital need of conferring legitimacy upon projects that might be otherwise forced “into the outskirts of the peer-control process, a cornerstone of academic excellence.”¹⁸ But they do not address the problem of publication itself. Many, if not most, scholarly digital editions have been published only in the sense that they have been made publicly available; they may be hosted and backed by respectable research institutions, and they may be the subjects of reviews in publications such as *RIDE*, but they have not been subjected to the same rigorous pre-publication review that is a matter of course for most traditional academic publications.¹⁹ The *LDLT* addresses the problem by replacing the traditional role of a commercial publisher with a partnership between the *DLL* and its affiliated scholarly organizations.²⁰

In a traditional publication process, commercial presses benefit from the labor of scholars. A scholar submits the product of her work to a press. If the product fits the profile of the press, an editor at the press consults other scholars about the submission. If they recommend publication, the press works with the original scholar to prepare the final draft for publication. In some instances, the press will ask the scholar or her institution to subsidize the publication with a subvention. After submission of the final draft, the press assembles the product and handles marketing and distribution efforts. Libraries purchase the product, and some scholars purchase copies for themselves. Meanwhile, editors of journals receive review copies and send them to other scholars, who write reviews that are published a few months later, possibly leading to more sales of the volume. When it is time to evaluate the original scholar’s performance, her colleagues at her home institution assess the product of her efforts, usually with the help of external reviewers, each of whom receives a copy of the volume.

Sometimes scholars at various stages of this chain receive remuneration for their efforts, but often this work simply falls under the category of professional

18 <http://ride.i-d-e.de/about/editorial/>, accessed 23.01.2018.

19 See Ducourtieux (2004) for a discussion of the distinction between publishing and making something publicly available.

20 In this way, the *DLL* aims to be the third party sought by Del Turco (2016) 234–235: “We need a reliable third party, such as universities and other research institutions, offering support and preparing an adequate infrastructure for long-term publishing of select digital editions.”

service. The press, however, is likely to be concerned with the bottom line as much as—if not more than—the quality of the scholarship. The cramped and abbreviated format of the apparatus criticus in traditional printed editions testifies to the commercial concerns of publishers. Like it or not, to publish or not to publish is ultimately a business decision. The point of this digression is to demonstrate that scholars do the bulk of the intellectual work in the publication process, but the publisher benefits from the sales.

The *DLL* aims to offer an alternative to this commercial process through its affiliation with learned societies. The idea is to leverage the infrastructure that learned societies already have in the form of committees or boards that oversee publications of various kinds.²¹ Often they enter into relationships with commercial publishers to handle production, marketing, and distribution, but the learned societies and their members are responsible for most of the intellectual effort: they establish missions and goals for their publications; they recruit editors for those publications; they rely on members to serve as reviewers, both pre- and post-publication. In short, they set and uphold the standards of scholarship, so their imprimatur on a publication should be sufficient to inspire confidence in readers of all types, including tenure and promotion committees.

According to this model, scholars have control over every aspect of publication, *ab ovo usque ad mala*. Scholars produce editions. Their peers in their professional communities set standards for publishing those editions and subject new proposals to peer review according to those standards, giving learned societies the confidence to place their imprimatur on editions. Scholars at the *DLL* develop tools and platforms for storing, retrieving, searching, and visualizing the data. And since the published data is openly available, other scholars may reuse the data in other projects without having to pay a licensing fee to a commercial publisher.

6 The Traveling Imprimatur

To protect the investment of the time, resources, and reputation of editors and the learned societies that publish *LDLT* editions, the *DLL* is implementing the “traveling imprimatur,” a concept developed by Jeffrey Witt, a member of the *DLL*’s advisory board.²² The traveling imprimatur takes advantage of the fact that a published *LDLT* edition is actually a specific version of a file in a version-controlled repository.

21 McGann (2013) 287 laments that digital scholarship “lacks the professional infrastructure that the scholarly book possesses by virtue of the mature social networks in which it is located.”

22 Witt (2016).

The edition bears a tag that identifies it as having met the standards of the learned society that published it. The tag follows the edition wherever it might be reused, providing verification that it is the authorized version.

The fact that a learned society's imprimatur applies only to a specific version of the edition's file means that editors can continue to update the file to reflect new developments in scholarship without affecting the contents of the published version. The imprimatur will still be in place wherever the *LDLT* edition is in use, since it refers not to the most current version in the repository, but to the authorized, published version. At some point, editors may resubmit their editions for review to have the imprimatur updated. This would be equivalent to a new edition of the text, since the previous edition still bears the tag of having been authorized and published. The point is to encourage ongoing scholarship while also giving users confidence in the reliability and trustworthiness of the published edition.

7 Conclusion

There is no question that scholarly editions have been worthy investments of the time and resources of individual scholars and institutions over the past several centuries. Their capacity for influencing scholarship even after they have been superseded is a testament to their long-term value. There is no reason to value scholarly digital editions any differently, provided that they do all of the things we have come to expect of scholarly editions, and provided that they remain accessible and usable. Indeed, soon it may be time, if the time has not come already, to consider dropping that modifier altogether so that we can return to speaking of scholarly editions, period. But that hinges on whether we recognize textual and information scholarship as distinct, yet complementary, disciplines. If we continue to expect scholarly digital editions to reflect equally high achievement in the fields of philology, human-computer interaction, data visualization, and information management, then they will retain their boutique nature, and they will appeal only to a niche market. Moreover, projects will continue to be expensive in terms of time and resources. But if we begin to value those different elements as important contributions to scholarship in their own rights, we will create a broader marketplace for ideas. The *Digital Latin Library* offers a model for letting scholars make these decisions individually and collectively, through their learned organizations. Whether the *DLL* will succeed remains to be seen, but it is safe to say that scholarly digital editions will be a good investment for scholars and institutions, as long as there are reliable outlets for peer review, publication, and dissemination.