Samuel J. Huskey

The Digital Latin Library: Cataloging and Publishing Critical Editions of Latin Texts

Abstract: The Digital Latin Library has a two-fold mission: 1) to publish and curate critical editions of Latin texts, of all types, from all eras; 2) to facilitate the finding and, where openly available and accessible online, the reading of all texts written in Latin. At first glance, it may appear that the two parts of the mission are actually two different missions, or even two different projects altogether. On the one hand, the DLL seeks to be a publisher of new critical editions, an endeavor that involves establishing guidelines, standards for peer review, workflows for production and distribution, and a variety of other tasks. On the other hand, the DLL seeks to catalog existing editions and to provide a tool for finding and reading them, an effort that involves the skills, techniques, and expertise of library and information science. But we speak of a "two-fold mission" because both parts serve the common goal of enriching and enhancing access to Latin texts, and they use the methods and practices of data science to accomplish that goal. This chapter will discuss how the DLL's cataloging and publishing activities complement each other in the effort to build a comprehensive Linked Open Data resource for scholarly editions of Latin texts.

Introduction

Although Latin texts have been available in electronic form for decades, there has never been an open, comprehensive digital resource for scholarly editions of Latin texts of all eras. In the era before the World Wide Web, collections such as the Packard Humanities Institute's (PHI) Latin Texts, Perseus, or Cetedoc made collections of texts available on CD-ROM, but those collections were limited by era (e.g., PHI and Perseus covered only Classical Latin texts) or subject (e.g., Cetedoc covered Christian Latin texts).[1] Matters improved with the wide

1 Cetedoc (sometimes known erroneously as CETADOC) was originally developed by the Centre Traditio Litterarum Occidentalium (CTLO). The full name of the database was "Cetedoc Library of Christian Latin Texts."

Samuel J. Huskey, University of Oklahoma

https://doi.org/10.1515/9783110599572-003

adoption of networked computing, but for many years collections of Latin texts were limited to a particular era (e.g., Perseus[2]), behind a paywall (e.g., Cetedoc, which became part of Brepolis' Library of Latin Texts[3]), or offline (e.g., PHI, which did not publish its texts online until 2011[4]). Sites such as The Latin Library and Corpus Scriptorum Latinorum are more expansive, but they have not kept pace with developments in technology, and since it is not always clear what the source of their texts is, they are of limited use for scholarly purposes.[5]

The Open Greek and Latin Project, however, promises to publish millions of words of Greek and Latin from all eras, along with robust resources for analyzing and reading the texts. As of this writing they have made significant progress toward that goal. Aside from the scale, what separates the Open Greek and Latin project from others is the focus on creating an open scholarly resource, with rich, citable metadata on the sources for the texts. But even the Open Greek and Latin project has established a boundary of 600 CE, which means that much of Medieval and Neo-Latin will be excluded.

But one thing that all of these resources have in common is that they omit the features that distinguish scholarly critical editions. That is, their texts lack an editorial preface that explains the history of the text and its sources, a bibliography of previous scholarship on the text, a critical apparatus with variant readings and other useful information, or any of the other items necessary for serious study. Whether the omission is because of copyright restrictions, the technical difficulty of presenting the information in a digital format, or the needs of the site's intended readership, it means that, with some exceptions, scholars must still consult printed critical editions for certain kinds of information.[6]

That is not to say that existing digital collections are useless for scholarship. After all, the goal of the Open Greek and Latin Project is not to publish critical editions, but to increase the amount of human-readable and machine-actionable Greek and Latin available online, and it promises to be an invaluable resource for a wide range of scholarship, from traditional literary and historical studies to

2 http://www.perseus.tufts.edu (last access 2019.01.31). It should be noted that the Perseus Digital Library expanded its Latin holdings to include authors from later eras, but on a limited basis. Its latest version (https://scaife.perseus.org, last access 2019.01.31) promises to be more expansive in terms of both texts in its library and tools available for studying them.

3 http://www.brepolis.net (last access 2019.01.31).

4 http://latin.packhum.org (last access 2019.01.31).

5 Corpus Scriptorum Latinorum: A Digital Library of Latin Literature: http://forumromanum.org/literature/index.html (last access 2019.01.31); The Latin Library: https://thelatinlibrary.com (last access 2019.01.31).

6 Kiss (2009–2013) is a notable exception. The catalog edited by Franzini et al. (2016–) contains details on other resources, but truly critical editions on the internet are still rare.

the latest developments in natural language processing. Rather, the point of this brief survey has been to define the space that the Digital Latin Library (DLL) means to fill: the collection and publication of critical editions of Latin texts from all eras, and the materials associated with them.[7]

To accomplish its objective, the DLL has two main initiatives: the DLL Catalog and the Library of Digital Latin Texts (LDLT). The purpose of the former is to collect, catalog, and provide an interface for finding Latin texts that have been digitized or published in digital form. The purpose of the latter is to publish new, born-digital critical editions of Latin texts from all eras. The rest of this paper will discuss these two wings of the DLL and their complementary goal of supporting new work in Latin textual criticism.

The DLL Catalog

As with other elements of the DLL, the "D" stands for "Digital" in a number of different ways. First and foremost, all of the items in the DLL Catalog are digital in some respect, either as digitized versions of printed materials or as digital texts.[8] Second, the catalog itself is digital, built with and operating entirely on open source technology. Most people will use the DLL Catalog via the web interface, but the datasets will be serialized in JSON-LD and available for downloading and reuse, in keeping with the best practices known as Linked Open Data.[9] Third, owing to the abundance of materials and the limited resources of the DLL, leveraging digital technology to ingest, process, and publish data is essential. Accordingly, building applications to facilitate those tasks is part of the scholarly endeavor of the DLL Catalog.[10]

Another way in which the DLL Catalog is digital is in its use of data modeling. Taking a cue from the Perseus Catalog[11] and using concepts from the

7 The Digital Latin Library project has been funded by generous grants from the Andrew W. Mellon Foundation's Scholarly Communications division from 2012 to 2018, and by ongoing institutional support from the University of Oklahoma.

8 See Sahle (2016) for an extended discussion of the difference between "digitized" and "digital." In short, a digital scan of a book may be referred to as "digitized," but not "digital," since it merely represents an object that exists in a non-digital format. To qualify as "digital," an edition must have distinct characteristics that would cease to function outside of the digital realm.

9 The repository is available at https://github.com/DigitalLatin (last access 2019.01.31).

10 See https://github.com/DigitalLatin/dllcat-automation (last access 2019.01.31).

11 http://catalog.perseus.org (last access 2019.01.31).

Functional Requirements for Bibliographic Records (FRBR)[12] model as a basis and data gathered from user studies, June Abbas and her team of researchers from the University of Oklahoma's School of Library and Information Studies designed an information behavior model to accommodate the different kinds of data to be stored in the catalog and the different ways in which users would interact with that data.[13] The following sections describe the resulting information architecture of the catalog and how it seeks to cater to the needs identified in Abbas' user studies.

Authority records

Authority records for authors and works provide the foundation for the DLL Catalog's information architecture. Each author of a Latin work has an authority record that identifies that author unambiguously and provides supporting attestations from a variety of sources to confirm the identity. In most cases, several forms of the author's name are recorded, especially the authorized name, which is usually identical to the authorized name in a major research library such as the U.S. Library of Congress, the Bibliothèque nationale de France, the Deutsche Nationalbibliothek, or others. Alpha-numerical or numerical identifiers such as the Virtual International Authority File ID or the Canonical Text Services identifier are also recorded, along with details about relevant dates and places. The purpose of an author authority record is to provide a single point of reference for individual authors. That way, searches for "Vergil", "Virgil", or "Vergilius" lead to the same information. Additionally, the cataloging process is more successful when automated matching algorithms have access to variant name forms.

Similarly, authority records for works support the vital functions of the catalog. Since dozens, if not hundreds, of works are known simply as *Carmina*, *Historiae*, or simply *fragmentum*, to take just three examples, it is important to have a means of disambiguating them. Accordingly, each work has its own authority record, with an authorized form of the title and any variant titles, along with information about its place in any collections, its author(s), and any abbreviations or other identifiers commonly in use.

Different content types for digitized editions, digitized manuscripts, and digital texts are the DLL Catalog's architectural frame. These content types

12 https://www.ifla.org/publications/functional-requirements-for-bibliographic-records (last access 2019.01.31).

13 See Abbas et al. (2015) for information about the methods and outcomes of the user studies.

store metadata related to specific instances of texts, each one connected to its creator and work through an entity reference so as to be discoverable in a variety of searches. Each record also contains a link to the external resource where the item can be found.

Contents

As of this writing, the DLL team has added authority records for over three thousand authors and nearly five thousand works spanning the time period from the third century BCE to the twentieth century CE. Many of those records were culled from information in standard reference works (e.g., *Clavis Patrum Latinorum*) and dictionaries (e.g., *Oxford Latin Dictionary, Thesaurus Linguae Latinae*), but records are also added nearly every time a new collection is added to the catalog, which is one of the reasons why the quest to catalog all Latin authors and works will be asymptotic.

Several collections are at different stages of being added to the catalog. With regard to digital texts, all of the items in following collections have been processed and cataloged: Perseus, PHI, Digital Library of Late-antique Latin Texts, and Biblioteca Italiana. Items on the related sites Musisque Deoque and Poeti d'Italia are in process and will be added by the end of 2019. These sites were selected because the sources of their texts are clearly identified and the texts themselves are openly available. Collections of texts behind a paywall (e.g., the Loeb Classical Library and Brepolis) are also in process, but since freedom of access is a priority, they will be added to the catalog at a later date.

As for digitized editions, efforts have focused on cataloging items in the public domain at resources such as the HathiTrust Digital Library, the Internet Archive, and Google Books.[14] Two categories in particular have received the most attention: early editions (*editiones principes*) and items in Engelmann's magisterial survey of Latin texts published between 1700 and 1878, *Bibliotheca Scriptorum Classicorum*. As of this writing, eighty-two early editions have been cataloged, including fifty-four editions of Latin texts published by Aldus Manutius. Over time, editions published by other early printers (e.g., Sweynheym and Pannartz, Jodocus Badius Ascensius), will be added to the collection. The survey of Engelmman's bibliography has so far yielded nearly three thousand

14 HathiTrust Digital Library: https://www.hathitrust.org (last access 2019.01.31); Internet Archive: https://archive.org (last access 2019.01.31); Google Books: https://books.google.com (last access 2019.01.31).

individual editions. Overlapping some of those are the records added in the effort to catalog all editions published in the history of the B.G. Teubner publishing house. To date, there are nearly nine hundred records in that collection.

The DLL Catalog also has a content type for manuscripts. Based on the guidelines of the Text Encoding Initiative's module for manuscript description, this content type is designed to be a resource for those wishing to find and view digital images of manuscripts of Latin texts. Since access to images of manuscripts varies widely among repositories, and since the metadata for manuscripts can be complex, progress on this initiative has been slower, but the catalog currently has nearly 1,300 records in process.

In sum, the DLL Catalog contained over 10,000 items when it was launched in the fall of 2018. Efforts to augment the catalog with items from other collections and library will be part of the DLL's ongoing mission to facilitate access to manuscripts, previously published critical editions, and other materials necessary for scholarly study of Latin texts.

The Library of Digital Latin Texts

Just as the DLL Catalog focuses on collecting historical editions and manuscripts of Latin texts, the Library of Digital Latin Texts (LDLT) focuses on publishing new, born-digital critical editions of Latin texts from all eras. The rest of this chapter will discuss what that means.

The subject of publishing digital scholarly editions is awash in paradoxes, some of them real, others only perceived. It is commonly assumed that younger scholars have an affinity for technology but pursue traditional modes of publication out of concern for the advancement of their careers. Conversely, it is assumed that senior scholars have more latitude for experimenting with new forms of publication, but lack the motivation or ability to learn new technologies. In both cases, the assumptions are only partly true. Although younger scholars are well-advised to publish their work in established outlets, it is not true that their age gives them any special facility with technology. Similarly, more established scholars do have some room for experimenting with publication formats, but it is ageist to assume that they necessarily have a block with respect to technology.

Leaving aside the false dichtomies of age and acumen, the LDLT aims to address the two real factors underlying those concerns. First, peer-review is essential to scholarly publications, so it is vital to have policies and procedures in place to ensure that LDLT editions meet the highest standards of the profession in that regard. Second, the digital format of the LDLT distinguishes it from

traditional critical editions in print, so it is important to take advantage of computing technology; at the same time, it is crucial not to exclude scholars from working on LDLT editions for lack of technical skill. The DLL has launched two initiatives to address both of these concerns.

Policies and procedures

It is one thing to publish something online in the sense of making it publicly available; it is something else entirely to submit one's work to review and criticism by one's peers in the field as part of an independent organization's publication process. Accordingly, the DLL publishes the LDLT through its affiliation with the Society for Classical Studies (SCS), the Medieval Academy of America (MAA), and the Renaissance Society of America (RSA). Throughout the planning and implementation stages, the DLL has convened regular meetings of an advisory board composed of representatives from all three organizations. The chief goal of these meetings was to devise and agree upon policies and procedures for subjecting LDLT editions to the same level of peer-review that other publications typically receive.

Since all of the organizations publish monographs or other print publications, they have the organizational structures in place for managing the process of receiving submissions, identifying potential reviewers, making final decisions to publish or not to publish the material, and working with a press to see the project through to completion. Submissions to the LDLT are handled in the same manner. First, scholars submit proposals for LDLT editions to the publications board of the appropriate organization. Depending on the organization and the nature of the text, the proposal may include, for example, the argument for the edition, a sample of the work, a description of the strategy and timeline for completing the edition, and a statement of the editor's qualifications. Second, the board reviews the proposal, with consultation of qualified peer reviewers, if necessary, and decides whether or not to pursue it. If the outcome is favorable, the proposal is entered into a database of projects, and the organization authorizes the DLL to begin working with the editor. If the final version receives a favorable recommendation from the board, the edition is published in a version-controlled repository under the control of the DLL.

Another part of this initiative is the drafting of publishing agreements between 1) the DLL and the affiliated learned societies, and 2) the editors of LDLT editions, the DLL, and the learned society under whose imprimatur the edition will be published. These agreements state the rights and responsibilities of all parties, especially with regard to the open license under which LDLT editions

are published. As of this writing, the Office of Legal Counsel at the University of Oklahoma, the DLL's host institution, is working with the DLL and the learned societies to finalize the agreements ahead of the publication of any editions.

Digital publication

Leveraging the digital nature of the LDLT means not only continuing to pursue and develop new methods for the use of technology with Latin texts, but also facilitating the participation of editors and other users of the LDLT who have varying levels of comfort with technology.

Key to this effort is clarifying what is meant by "digitial scholarly edition" in the first place, at least within the confines of the LDLT, since that term is in use elsewhere for everything from simple HTML documents to complex, multimedia databases. Indeed, a quick survey of the editions cataloged by Franzini et al. (2016–) reveals just how capacious the usage of "digital scholarly edition" is. As of this writing, the catalog has two hundred ninety-six items in general. Application of the filters for "scholarly", "digital", and "edition" reduces that number to two hundred thirty-seven. Those filters are based on the work of Sahle, who offers a useful way of thinking about the digital component (2016, 28): "Scholarly digital editions are scholarly editions that are guided by a digital paradigm in their theory, method and practice." But his discussion reveals that the "digital paradigm" is closely bound to presentational format. That is, by his definition, editors of scholarly digital editions are accountable for the quality of not only their textual scholarship, but also the design, implementation, and functionality of the interface and its accompanying technology. Although it is certainly the case that arguments about a text can be advanced through information visualization, the DLL asserts that human-computer interaction, data visualization, and user interface design should be taken seriously as scholarly disciplines unto themselves. Moreover, although some textual scholars might have the aptitude and capacity for developing mastery of these additional disciplines, they are the exceptions.

Accordingly, the LDLT aims to separate content from presentational format as much as possible. The qualification "as much as possible" is a nod to the fact that any representation of textual data, whether in plain text, encoded in Extensible Markup Language (XML), or on paper has a presentational format that influences how a reader (human or machine) interacts with it. Nevertheless, since data visualization and interface design add several layers of complexity to the traditional task of editing a text, an LDLT edition consists of

the contents of a single XML file published in a version-controlled repository. As will be explained below, the DLL provides some official and experimental visualizations of the data in an LDLT edition as part of its ongoing scholarly research initiatives. Additionally, since LDLT editions are published on an open basis, anyone is free to reuse the data for other projects, including, but not limited to, the design and implementation of independent reading environments and data visualizations. But the edition file itself includes only prefatory materials, text, and scholarly apparatus; there are provisions for including expanded notes on the text, but extended commentary is outside of the scope of an LDLT edition. Additionally, editors are encouraged to include research notes, images, transcriptions of manuscripts, collation tables, and other materials in the repository that contains the edition file, for the sake of users who wish to conduct further research or who might have other uses for the research data. The option to include such materials is also in recognition of the scholarly approach that holds that a text's multiple versions in its various sources cannot be adequately conveyed to readers in a single critical edition.[15]

If the DLL left all decisions about content, encoding strategies, and presentational formats to editors, the LDLT would be just a loose collection of projects, each with its own unique approach and features, and it would be viable as a publication forum only for editors with the requisite technical skill. Although prescribing the encoding method and separating content from presentation does set some limits on what may be included in an LDLT edition, it also ensures that LDLT texts will have features in common, which means that they will be more useful as a uniform corpus of texts. It also means that they will work with the LDLT's applications.

Just as the "D" in the DLL Catalog includes the development of digital tools for processing information for the catalog, the "D" in LDLT emcompasses the digital tools and methods developed by the DLL for facilitating the creation and use of digital editions. The following tools and methods are the DLL's independent scholarly research outcomes in support of the LDLT project.

15 Such is the prevailing view of the essays collected by Apollon et al. (2014). See also Heslin (2016), who considers textual ctiticism as a "mental disorder," and who argues in favor of variorum editions instead. During a panel discussion at the 2018 annual meeting of the Society for Classical Studies, Heslin appeared to agree that the LDLT's approach of providing a canonical edition and access to transcriptions and collation materials is a good way of bridging the divide between new and traditional philology.

Encoding guidelines

Huskey and Cayless' "Guidelines for Encoding Critical Editions for the Library of Digital Latin Texts" are the foundation for the other research projects associated with the LDLT. A customization of the Text Encoding Initiative's guidelines, with strong ties to Epidoc, the LDLT's encoding guidelines provide instructions for using XML to represent the various kinds of information typically found in critical editions, including the preface, main text, the various types of scholarly apparatus, and ancillary materials.

The majority of the work in developing the guidelines involved manually encoding a model edition. Giarratano's first edition of the bucolic poetry of Calpurnius Siculus was selected for this project for several reasons. First, Calpurnius Siculus' seven *Eclogues* add up to about the length of a "book" of Classical Latin poetry or prose: 759 lines of poetry. That seemed to be a manageable and reasonable size for a model text. Second, the textual tradition involves a number of interesting problems, including lacunae and the transposition of words, lines, and whole stanzas. Third, Calpurnius' poetry has attracted the attention of many illustrious figures in the history of philology, including Boccaccio, Heinsius, Burman, Scaliger, and Wilamowitz, among others, so the bibliography is rich and interesting from a historical point of view. Finally, Giarratano's edition features an ample and detailed apparatus criticus, with plenty of edge cases for testing the limits of the data model. In consultation with Cayless, Robert Kaster, and Cynthia Damon on technical and textual matters, and with the assistance of several students at the University of Oklahoma,[16] Huskey encoded every line of poetry and every entry in the apparatus criticus, along with the preface, description of manuscripts, and the conspectus siglorum.[17] At the same time, Huskey and Cayless collaborated on compiling the encoding patterns, rules, and techniques into a document that eventually became the guidelines.

To test the applicability of the guidelines to other kinds of texts, the DLL enlisted some scholars to prepare pilot editions for the LDLT. Whether or not these editions will be published is up to the learned societies affiliated with the project to decide, but having materials for testing purposes has been invaluable. To ensure broad applicability of the guidelines, we selected a variety of texts, including books 9–12 of Servius' commentary on the *Aeneid* (edited by

16 Shejuti Silvia, Bharathi Asokarajan, Sudarshan Vengala, Vamshi Sunchu, Alexandra Owens, and Matthew Mitchell.

17 The current version of the model edition of Calpurnius Siculus' bucolic poetry may be found at https://github.com/sjhuskey/Calpurnius_Siculus (last access 2019.01.31).

Robert Kaster), Pseudo-Caesar's *Bellum Alexandrinum* (edited by Cynthia Damon), Peter Plaoul's Commentary on the *Sentences* of Peter Lombard (edited by Jeffrey Witt), and the Book of Genesis from the Codex Amiatinus (edited by Andrew Dunning). The encoding of each text has contributed to the evolution of the guidelines and preliminary results indicate that they will accommodate the majority of texts submitted to the LDLT.

In addition to providing uniform guidelines for producing material for the LDLT, the guidelines themselves are also a plank in the DLL's platform for promoting different forms of digital scholarship. More than just an application of existing instructions for encoding data, the guidelines are an argument about the form and function of critical editions. The addition of each new text to the LDLT will test that argument, and the guidelines will evolve to accommodate previously unforeseen scenarios.

Automated encoding

Editors have the option of encoding their editions themselves, using any of the many commercial and open source products for writing and editing XML, but they can also avail themselves of the automated encoding processes developed by the DLL. These automated processes have been developed in part as a way of testing the validity of the LDLT's data model. The argument is that if the encoding guidelines provide a sufficiently detailed structure for the various kinds of textual data, it should be possible to automate much of the standard encoding processes through algorithms based on the guidelines. For example, Felkner and Huskey have developed a series of Python scripts that automate the encoding of nearly all of a prospective LDLT edition, freeing editors to focus on textual matters instead of low-level encoding issues that do not require editorial scrutiny. Anything that cannot be encoded automatically is likely to require the editor's input regarding the precise nature of textual data in question, effectively highlighting the fundamental role that human judgment continues to play in textual criticism. Whether editors resolve those issues independently or in consultation with the DLL, the outcome is likely to influence further development of the automated encoding tools, and possibly the guidelines themselves.

LDLT viewer

The LDLT viewer, designed by Hugh Cayless, provides much of the functionality that June Abbas, co-PI on the DLL project, identified as necessary or desirable

through her user studies. Based on the CETEIcean reader Cayless developed for the Text Encoding Initiative,[18] the LDLT viewer leverages HTML5 Custom Elements to avoid the need to process the XML data before displaying it in an internet browser. Instead of requiring the intermediate step of a data transformation via XSLT or some other method, the LDLT viewer application renames the elements in accordance with Custom Elements conventions. The resulting HTML preserves the structure of the original XML file, but it renders the data in a way that is more friendly to human readers.

The LDLT viewer preserves the traditional layout of a critical edition, with the text occupying the main portion of the display and the apparatus criticus appearing at the bottom of the screen, but there are also some important innovations. Chief among them is the additional dynamic apparatus display. Clickable icons appear to the right of any portion of the text that has corresponding data in the apparatus. Hovering the mouse over an icon causes the lemma in question to be highlighted in the main text. Clicking on the icon activates a dialog box that reveals the apparatus data related to that lemma. Clicking on a variant reading causes the variant to be substituted for the lemma in the main text so that it can be evaluated *in situ*. It will also cause related variants to be substituted simultaneously. For example, if a manuscript has two words or phrases transposed and that tranposition has been encoded in sufficient detail, clicking on one word or phrase will activate the other one, too, lest the viewer display a version of the text that does not exist in some source.

If an editor has tagged variant readings with terms from the taxonomy of variants included in the LDLT's encoding guidelines, other functionality is also enabled in the form of filters.[19] Users who do not wish to see apparatus entries concerned solely with orthographical variants can activate a filter to hide variants with that tag. Similar filters are available for morphological and lexical variants. A reset button restores the edition to its original state.

During development of the LDLT viewer, the DLL pooled some resources with the Open Philology Project to support the development and expansion of the Alpheios Reading Tools into Javascript libraries that can be deployed independently of specific browsers.[20] The LDLT viewer implements the Latin word parser and dictionary lookup libraries so that users can click on words to see automated lexical and morphological analyses.

18 https://github.com/TEIC/CETEIcean (last access 2019.01.31).

19 The section "Tagging Readings for Analysis" (https://digitallatin.github.io/guidelines/LDLT-Guidelines.html#apparatus-criticus-analysis, last access 2019.01.31) is the result of a collaboration between Huskey and Robert Kaster.

20 https://alpheios.net (last access 2019.01.31).

The LDLT viewer also operates on a framework compatible with Canonical Text Services, which means that users will be able to use CTS URNs to cite specific passages in texts.

Data visualization

To demonstrate the potential applications of data visualization scholarship to Latin texts, the DLL is making available a downloadable desktop application pre-loaded with a number of visualizations developed by Chris Weaver, another co-PI on the DLL project, and his students. Using the framework from his Improvise visualization application,[21] Weaver and his students have developed techniques to represent textual data in ways that will highlight the potential uses of visual data analysis for Latin textual studies. These visualizations are the most experimental of the the DLL's projects. Consequently, they should be considered candidates for further development after their initial release.

VariantFlow, developed by Shejuti Silvia, is a storyline visualization that represents manuscripts and other sources for a critical edition as individual lines that tell the story of variation in the text through their intersections and divergences. Proceeding along a horizontal plain that tracks the "story" of the critical apparatus from left to right, the lemmata serve as checkpoints. Observing how the storylines of the sources converge or separate throughout the text's overall "story" can provide a new perspective on the textual tradition.

TexTile, developed by Bharathi Asokarajan, uses pixels to represent the sources of a text and colors to indicate the degree of variance from the lemma, based on a string metric known as Levenshtein distance.[22] This tool presents the data in three different levels of focus: individual apparatus entry, line, and text chunk. Users control which level they see with a slider that brings different lemmata into focus. A spectrum of colors represents the degree to which a source varies from a given lemma or an edition's text in general.

Encodex, developed by Weaver, is a visual interface that integrates a regular text viewer with the visualizations mentioned above. As users scroll through the text, various kinds of highlighting will alert them to words or phrases with corresponding data in the apparatus. The other visualizations

21 http://www.cs.ou.edu/~weaver/improvise/index.html (last access 2019.01.31).

22 For more on Levenshtein distance, see https://en.wikipedia.org/wiki/Levenshtein_distance (last access 2019.01.31).

are synchronized with the scrolling operation in Encodex, giving readers a dynamic environment in which to explore different ways of looking at the text.

Conclusion

In some respects, this chapter has been about the future, since the various components described above will have their official launch while this book is in press. But even after their launch, there will never be a point at which the Digital Latin Library can be said to be complete. Although the DLL Catalog will launch with a large number of authority records and individual items, scouring the corners of the internet for Latin texts will be an ongoing project, both in terms of cataloging the content and developing new tools and methods for using it. Similarly, building the LDLT will be a long-term project, considering the number of Classical, Medieval, and Neo-Latin texts in need of new treatment as digital editions. But it is worth doing, especially if the availability of a sustainable outlet for publishing high quality, peer-reviewed Latin texts on an open basis encourages a new generation of scholars to continue the tradition of textual criticism.

Bibliography

Abbas, J.M.; Baker, S.R.; Huskey, S.J.; Weaver, C. (2015): "Digital Latin Library: Information Work Practices of Classics Scholars, Graduate Students, and Teachers". In: Proceedings of the Annual Meeting of the Association for Information Science and Technology. Silver Spring, MD: Association for Information Science and Technology. https://www.asist.org/files/meetings/am15/proceedings/openpage15.html (last access 2019.01.31).

Apollon, D.; Bélisle, C.; Régnier, P. (eds.) (2014): Digital Critical Editions. Urbana, Chicago, and Springfield: University of Illinois Press.

Asokarajan, B.; Etemadpour, R.; Huskey, S.J.; Abbas, J.M.; Weaver, C. (2016): "Visualization of Latin Textual Variants using a Pixel-Based Text Analysis Tool". In: Proceedings of the International Workshop on Visual Analytics. Geneva, Switzerland: The Eurographics Association. http://diglib.eg.org/handle/10.2312/eurova20161119 (last access 2019.01.31).

Crane, G.R.; Berti, M.; Geßner, A.; Munson, M.; Selle, T.: The Open Greek and Latin Project. http://www.dh.uni-leipzig.de/wo/projects/open-greek-and-latin-project (last access 2019.01.31).

Elliott, T.; Bodard, G.; Cayless, H. (2006–2017): EpiDoc: Epigraphic Documents in TEI XML. http://epidoc.sf.net (last access 2019.01.31).

Engelmann, W.; Preuss, E. (1882): Bibliotheca Scriptorum Classicorum. Volume 2: "Scriptores Latini". Leipzig: Wilhelm Engelmann.
Felkner, V.K.; Huskey, S.J.: "Digital Latin Library: Automation". https://github.com/DigitalLatin/automation (last access 2019.01.31).
Franzini, G.; Andorfer, P.; Zaytseva, K. (2016–): Catalogue of Digital Editions: The Web Application. https://dig-ed-cat.acdh.oeaw.ac.at (last access 2019.01.31).
Giarratano, C. (1910): Calpurnii et Nemesiani Bucolica. Naples: Detken et Rocholl.
Heslin, P. (2016): "The Dream of a Universal Variorum: Digitizing the Commentary Tradition". In: C.S. Kraus; C. Stray (eds.): Classical Commentaries: Explorations in a Scholarly Genre. Oxford: Oxford University Press, 494–511.
Kiss, D. (2009–2013): Catullus Online: An Online Repertory of Conjectures on Catullus. http://www.catullusonline.org (last access 2019.01.31).
Sahle, P. (2016): "What is a Scholarly Digital Edition?". In: M.J. Driscoll; E. Pierazzo (eds.): Digital Scholarly Editing. Cambridge: Open Book Publishers, 19–39.
Shejuti, S.; Etemadpour, R.; Huskey, S.J.; Abbas, J.M.; Weaver, C. (2016): "Visualizing Variation in Classical Text with Force Directed Storylines". In: Proceedings of the Workshop on Visualization for the Digital Humanities. Baltimore, MD: IEEE.