IDEAH • Vol. 1, Iss. 1 (DHSI 2017 & 2018)

Mining Digital Humanities Archives to Produce Accessible Material Editions of Scholarly Texts

Gia Alexander

Published on: May 31, 2020

DOI: 10.21428/f1f23564.4e680428

License: [Creative Commons Attribution 4.0 International License \(CC-BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

Rationale

Scholarly practices within book history, textual studies, and literary criticism continue to emphasize close interaction with the material text, even amidst growing technical advances in and academic acceptance of digital scholarship (Chevarie et al. 210-20). In the introductory chapter to his 2019 monograph, *How and Why Books Matter: Essays on the Social Function of Iconic Texts*, James Watts emphasizes the sustained value of the physical text, arguing for its continued cultural relevance as ritual icon (7-29). Analytical, or physical, bibliography necessitates physical interaction with the material text to understand and describe how books are constructed (Gaskell 313-20). Textual studies, or scholarly textual editing, focuses on the presentation of the material text on the page, identifying and analyzing differences between separate documents as individual instantiations of a text (Shillingsburg ix). Textual scholars often then produce rigorous new editions of the documents they examine to present how a text has changed over time through successive previous editions and printings (Kelemen and Reiman 3-28). The Society for the History of Authorship, Reading, and Publishing (SHARP), the premiere professional society for the history of the book, emphasizes the material text as the foundation of book history as a discipline (SHARP) and its pedagogy (Hawkins 15; Shep 38-43). Literary criticism demands close reading of text and theoretical interpretation of its nuances. Such analyses often produce wide variations of interpretation with stratified degrees of rigor and credibility; even the placement of a comma can make a difference.

With so much emphasis on the material text, this piece may, at first glance, seem out of place in an electronically published collection of essays about digital humanities (DH). And yet, the relationship between physical and digital text is indeed symbiotic, with one regularly being used to preserve and augment the other. Indeed, in a 2015 lecture, Jerome McGann stated that, “[w]e will likely spend the next 50 years digitizing our culture.” DH has proven itself all but indispensable over the last two decades in capturing, curating, archiving, and preserving human creative output and making that output available to wider audiences who are not able to travel to the libraries, museums, and archives where the physical artifacts are kept after digitization. However, the digitization of text does not necessarily make it accessible to those audiences who have print disabilities, defined here as physical or cognitive conditions that limit or prohibit one’s ability to access, read, and analyze written or printed material. Such access barriers particularly affect those who need to interact physically with text, such as librarians, special collections curators, book historians, textual

scholars, or literary critics. Providing means of accessibility to physical and digital text not only broadens the audience exposure for any given material document, physical collection, or digital archive, but also opens the related fields of study to more diverse scholars, including those with print disabilities.

In addition to critical considerations of ethics, audience, and inclusivity, there is also a legal mandate for federally funded projects and those housed at public universities in the United States to adhere to accessibility best practices, and DH projects are included in such a mandate.

Hence, DH accessibility is currently a key area of emphasis within the discipline. In 2014, for example, the National Endowment for the Humanities sponsored a series of four two-day workshops around the US called “Building an Accessible Future for the Humanities.” These round-table events brought together both DH scholars and stakeholders to discuss “theoretical and practical approaches for making digital humanities scholarship accessible to blind, low-vision, deaf, and hard-of-hearing users.” (NEH). Each year, the Digital Humanities Summer Institute (DHSI), held at the University of Victoria, British Columbia, Canada, offers a week-long, intensive course on accessibility in digital environments (Williams and Templeton). National and international DH conferences, such as the annual Chicago Colloquium on Digital Humanities and Computer Science, continually invite discourse on accessibility (CCDHCS18).

Digitizing text can certainly go a long way toward making it accessible to users who have print disabilities. Once a document is in digital form, usually accomplished either via coding or optical character recognition (OCR), users can enlarge the text using screen magnification software or port it to audio by way of text-to-speech output. Several commercial software packages are available for this purpose, including Job Access with Speech (JAWS)¹ and ZoomText, as well as free applications such as No Vision Desktop Access (NVDA) and Thunder. More recently, mainstream applications for producing and consuming text, such as Microsoft Office and Adobe Acrobat, have incorporated read-aloud capability. Built-in smartphone applications, such as VoiceOver on the Apple/iOS platform and VoiceAssistant on the Android platform, provide accessible text-to-speech capabilities on the go.

Access barriers remain, however, in that neither magnification nor audio output entirely preserves many of the structural or physical aspects of the material text or the intellectual and cognitive benefits of physical reading over audio, benefits such as improved comprehension and retention of information. Indeed, in their 2012 article,

“Recognition Memory for Braille or Spoken Words: An fMRI Study in Early Blind,” Burton, Sinclair, and Agato conclude that Braille readers have better memories, and that they learn and retain new words more easily than those who rely on listening alone (22). Subjects, all blind at or near birth, were divided into two groups; one group read Braille and the other listened to audio. Each participant entered an fMRI machine and was presented with “old” words that they knew well and “new” words they had just learned from their reading exercise. The scientists studied the level of oxygen used by each person’s brain during word presentations as evidenced on each participant’s fMRI. Those who listened to the “new” words showed greater oxygen use, and thus greater effort, in learning and recalling the words, whereas those who read the material in Braille showed greater facility with less effort in recall (Burton et al. 33-4). In a similar 2013 study titled “Differential Cognitive and Perceptual Correlates of Print Reading Versus Braille Reading,” Veispak, Boets and Ghesquiere compared sighted print readers with blind Braille readers to measure “reading, auditory, speech, phonological and tactile spatial processing” (372). These authors conclude that Braille reading involves different brain functions than listening. They go on to assert that Braille readers have superior retention skills when they do listen than those of their sighted counterparts who read print. This work indicates that the use of Braille not only improves listening skills in blind users, but also improves comprehension overall (378-85).

Reverse-engineering the means by which text is digitized can capture important enhancements to accessibility made during the digitization process, namely Text Encoding Initiative (TEI) markup, and then render a physical copy that enables and preserves scholarly interaction with the material text. For those readers who may be newcomers to DH, TEI, the encoding method digital humanists use to prepare text for digital archiving, designates markup tags not entirely dissimilar from HTML codes for websites, such as those used to denote headings or paragraph breaks. TEI, however, offers hundreds of coding tags that enable editors to denote everything from line and stanza breaks in poetry to the color and type of ink used on a handwritten letter to a hole in the material on which a given text is inscribed. As an extended example, not only can TEI encode a hole in the paper, but it also allows for coded description of the nature of said hole: has the paper been punctured, torn, burned, or chewed by an animal, or, in the case of parchment or vellum, does the hole appear to be the result of overstretching the hide in the harrow (stretching frame) while being prepared? TEI also provides codes for emendations, such as redaction, insertion, marginalia, strike-through, or, in the case of the hole, method of repair, if any. For instance, holes or

tears in parchment or vellum would often be sewn closed. Ultimately, TEI provides for the long-term preservation of digitized text along with all of the descriptive tags in that it is open-access, and its output is not platform-dependent (TEI).

As a literary scholar and Victorianist book historian with lifelong low vision, I have had to work through several barriers to access material text. In fact, want of such access motivated me to study DH and gain expertise in accessibility in digital environments. In 2014, I embarked on a DH project to make both print and digital scholarly texts, such as critical editions, more available and accessible. The following year I published a paper (Alexander) in which I set forth the theory behind and possibility of combining TEI-encoded digital text with uncontracted Braille output to produce accessible scholarly editions that enable physical interaction with material text. In 2016, I received funding to purchase JAWS software, a Refreshable Braille Display (RBD, essentially a Braille keyboard that has a row of pins that move up and down to create Braille characters), and a Braille printer to bring the project to proof-of-concept. Having successfully accomplished this stage in 2017, I presented this work, including hard-copy, TEI-encoded, uncontracted Braille output, at DHSI 2018. The present paper now serves as the other bookend to the project. The next section reviews the most salient theoretical and practical points of accessibility in digital environments. While the work proposed in my initial paper began by hand-coding with TEI markup the introductory lines of the Prologue to Geoffrey Chaucer's *The Canterbury Tales* as a demonstration case, the latter sections of the present paper explain how to locate and acquire existing TEI-encoded files of vetted, scholarly text to produce accessible physical editions. This technical process of using digitized archival material, however, must first be grounded in the theories of accessible DH to clarify the logical trajectory of this approach.

Accessibility and DH

George Williams and Erin Templeton begin their annual DHSI workshop on accessibility in digital environments with a long view, urging participants to take heed of Stephen Covey's second habit to "begin with the end in mind." By this reference, I mean that Williams and Templeton emphasize that accessibility should be planned for and built into DH projects at their inception, thus avoiding later retrofit. Therefore, this workshop begins with group discussion and definition of the overarching concepts of disability, accessibility, accommodation, and universal design (Williams and Templeton). Within the context of DH, a disability is a barrier that prevents an audience from accessing and benefitting from the content of a project. As a discipline,

DH largely ascribes to the social model of disability, the notion that a disability is a social construct that excludes people who, because of a physical, mental, cognitive, or emotional condition, are unable to fully participate in the social activity at hand (Barnes 65-6). In this theory, the problem—the disability—is the socially-constructed barrier, not the user's medical condition. For example, the social model of disability would fault a DH project including non-closed-captioned videos as disabling to users who are deaf or hard of hearing. The obverse is true of the medical model of disability, which situates the problem as the user's limiting medical condition, a circumstance that needs to be fixed or cured (Barnes 66). From the standpoint of the medical model, then, in the above example deafness would be the disability, and responsibility for overcoming this problem would lie with the medical community, not with the DH project containing the inaccessible videos. Because DH rejects the medical model, the responsibility for providing accessibility lies squarely within the planning and coding skills of the project developer. In DH, accessibility means anticipating and providing for users with different types of access needs (Williams and Templeton). In videos, for example, closed captioning would provide accessibility to users with hearing impairments, while a separate audio track or text transcript containing video description would provide accessibility to users whose access needs are related to their vision.

An accommodation, or Reasonable Accommodation under the Americans with Disabilities Act (ADA) (United States), is a modification made to an existing situation, or DH project in the present case, to remove a specific access barrier for a given user or group of users (Williams and Templeton). To carry forward the video example, adding description to an existing video in a DH project in response to a request for it from a blind user would constitute an accommodation. This addition would also constitute a retrofit, in that the project developers did not anticipate the need for descriptive video ahead of time, responding only when receiving the request for accommodation. Such retrofits are generally undesirable and can be especially problematic in legacy DH projects that, while still online, are no longer being actively maintained or updated. Another primary example of problematic retrofit is the lack of alternative text, coded "alt-txt," for images. Sometimes, even the most well-meaning developers never get the time or the funding to go back and retrofit an entire DH project or archives that may contain hundreds of digitized images. While software such as ImageReader, an available add-on to ZoomText, is currently being developed to provide some description for online images that are missing alternative text, the rendered descriptions still tend to be very general, such as "image of nature,"

“drawing of buildings,” “photograph of books,” etc. Such software certainly does not capture the DH project developer’s reason for curating the images in the first place, nor can it provide context, such as describing the role of one image within a series. For that, we still need a human, in this case the digital humanist.

Many digital humanists will strive for universal design in developing their projects in such a way as to allow for access by as wide an audience as possible, including users both with and without disabilities. (Williams and Templeton). Including a translation tool is an example of universal design not necessarily germane to disability or accessibility—unless, for instance, a blind user’s text-to-speech synthesizer algorithm does not support the translated language. For example, medievalists with visual access needs currently struggle with this very issue when trying to port Old or Middle English text to audio output. In “Universal Design and Its Discontents,” Chaucer scholar Rick Godden and his collaborator Jonathan Hsy raise caution concerning universal design, emphasizing that it can never be completely universal and warning that unnecessarily overzealous attempts at universality can lead to scope creep and problems with project sustainability. Hsy and Godden certainly advocate for building in as much accessibility as possible, but they also task DH developers with the responsibility to aim for the widest audience they can most reasonably expect considering the subject matter of their projects.

While DH accessibility writ large comprises text, sound, video, imagery, and even 3-D rendering of objects, the next sections of this work draw the focus back to the material text and show how physical text that becomes digital can then be re-rendered to produce an accessible material edition.

Accessibility and Text

If the myth of Homer’s blindness is to be believed, then the oral tradition that he used meant that some of humanity’s earliest works of literature were “born accessible,” at least to those ancient Greeks who were also blind. Indeed, even millennia later, audio output remains the typical first course of action to make text accessible to those who cannot see it. It follows, of course, that even the earliest written texts were accessible to those who could not hear well, but who could see and read. Inscriptions into stone and clay may well have been accessible to individuals who could feel and make sense of the inscribed symbols. Magnification lenses that became available in the West during the Renaissance may have been the first true assistive technology aimed at making text accessible to those with poor eyesight. The next major advancement in textual accessibility would not come until the nineteenth century with the development

of “night writing,” a process used by various militaries, including the Napoleonic armies, that created embossed text (raised letters) that could be felt in the dark, when light would give away one’s position and messages read aloud could be intercepted by eavesdropping spies. Night writing would inspire Louis Braille to create his system of raised dots in 1819 (Braille Institute).

As the means to create tactile print increased, the Bible became one of the first texts to be made specifically in an accessible format (Braille Institute). A later form of night writing known as “Boston Line Letter,” one of five raised-letter systems in use (Braille Institute) was used to create accessible Bibles in the late nineteenth century until Braille became the universal tactile format (Museum of the Bible). As James Watts points out in his treatment of the continued relevance of books, the Bible stands out as the preeminent iconic text. In the present work, it stands as a prime example of the continuing need, at least for many Christians, to interact with the physical text. To some believers, the Bible app, the eBible, or the Kindle Bible simply will not do. The Bible, as in the physical book, acts as a ritual talisman; the good Christian is supposed to be “Bibled;” that is, they are supposed to possess their own Bible and physically read that Bible, engaging in embodied interaction with the Word of their God (Howsam 2).

With the advent of analog recording at the end of the nineteenth century, audio became the dominant means of producing accessible text throughout the twentieth century. In 1931, the United States Library of Congress established the National Library Service for the Blind and Physically Handicapped (known colloquially as the NLS) (Library of Congress). Though no longer politically correct in name, this service began by collecting embossed and Braille books and making them available through the mail to patrons with print disabilities (Library of Congress). Beginning in 1935 and lasting well into the 1980s, live readers would read books aloud and the recordings, known as “Talking Books,” were pressed on vinyl discs similar to the 78 rpm record format (Library of Congress). Patrons would receive, free of charge, a talking-book playback machine (essentially a record player calibrated to the disk format) and borrowing privileges for the recordings (Library of Congress). Starting in the 1970s, recordings began to be produced on four-track audio cassette; for many years thereafter, patrons used both types of machines and received recordings in both formats, as it took time to remaster the original vinyl disks to cassette (Library of Congress). At the turn of the present century, the preparation and dissemination of Talking Books finally became digital via the Digital Audio Information System, or DAISY format. Patrons now receive recorded books on USB drives or download them

directly to their own USB-connected playback devices, smartphone, or tablets from the NLS Braille and Audio Reading Download (BARD) service (nlsbard.loc.gov) (Library of Congress). Today, it is actually possible to download from BARD and similar services digital Braille, or .BRF, files that can then be de-digitized back into physical form by porting them to an RBD or Braille printer. However, these .BRF files are not TEI-encoded, so they do not offer any of the structural orientation to the material text that a TEI-encoded, uncontracted Braille file does. An explanation of uncontracted Braille and why it is the best option for producing TEI-encoded material scholarly editions appears in this article's appendix.

In addition to BARD, other subscription services, such as Bookshare (bookshare.org), offer digital text, primarily via synthesized audio instead of a recording of a live reader, to patrons with print disabilities, whereas other services such as Audible, Kindle, and Nook offer digitized text to any user, disabled or not, for a nominal fee. Bookshare, however, has recently begun to offer not only DAISY audio, but also downloads in Rich Text Format (.rtf) Microsoft Word, and even epub (Kindle, Nook) formats. However, these formats still strip the structure and feel of the original material text. An additional drawback for scholars of literature is that most of these digitized recordings and downloads are created from trade or paperback editions for an audience of general, as in non-scholarly, readers. Granted, today primary and scholarly texts such as critical editions can be scanned via OCR, then rendered as plain text that can be enlarged or ported to audio. OCR, however, while much improved, remains not completely reliable, and is thus an unstable format for a textual scholar. The resulting output of this process also still lacks the detailed layout description of the material text that TEI encoding provides and the experience of embodied textual interaction that Braille provides. While a comprehensive discussion of the Braille writing system lies beyond the scope of the present work, the appendix of this work provides a brief introduction and explains the role of uncontracted Braille in the process I developed to create accessible, scholarly, material editions of text.

This research contributes to the field of digital humanities a means of providing access to digitized material that is visual in nature. DH developers can use the tools and techniques described in this work to expand their audiences and fulfill legal accessibility mandates for their federally funded projects. Improved accessibility to DH resources also further enhances the ethical and inclusive standing of the discipline. The specific processes described in this article and its appendix will enable users, particularly scholars with visual access needs, to work with the best available scholarly editions of texts in their respective subfields. Further, they will have the ability to

perceive and understand the physical structure of a text as described in its TEI code, and to interact materially with that text. Ultimately, this work opens the fields of literary and textual studies to interested scholars who have print disabilities.²

Works Cited

Alexander, Gia. "Chaucer's 'General Prologue' to The Canterbury Tales as a Case for Accessible Scholarly Editions Using TEI-Encoded Uncontracted Braille." Special Issue of *Journal of Interactive Technology and Pedagogy*, 8, 2015, n.p._
jitp.commons.gc.cuny.edu/chaucers-general-prologue-to-the-canterbury-tales-as-a-case-for-accessible-scholarly-editions-using-tei-encoded-uncontracted-braille/.

Braille Bibles International. "New King James Bible."
www.braillebibles.org/Braille.html, accessed March 28, 2014.

Braille Institute. "History." www.brailleinstitute.org/about-braille-institute/history, accessed April 9, 2014.

Barnes, C. "The Social Model of Disability: A Sociological Phenomenon Ignored by Sociologists." *The Disability Reader: Social Science Perspectives*, edited by Tom Shakespeare, Bloomsbury, 1998, pp. 65-78.

Burton, Harold, et al. "Recognition Memory for Braille or Spoken Words: An fMRI Study in Early Blind." *Brain Research*, no. 1438, 2012, pp. 22-34.

Chevarie, Joan F., et al. "Digital Scholarship in the University Tenure and Promotion Process: A Report on the Sixth Scholarly Communication Symposium at Georgetown University Library." *Journal of Scholarly Publishing*, vol. 40, no. 3, April 2009, pp. 219-230.

Chicago Colloquium on Digital Humanities and Computer Science. "DHCS 2018." 2018. ctsdh.org/dhcs2018/, accessed June 11, 2018.

Covey, Stephen R. *The Seven Habits of Highly Effective People: Powerful Lessons in Personal Change*. Simon & Schuster, 1989.

Gaskell, Philip. *A New Introduction to Bibliography*. Oxford UP, 1972.

Hawkins, Ann, editor. *Teaching Bibliography, Textual Criticism and Book History*. Routledge, 2006.

Howsam, Leslie. "Saints in Publishing." *Cheap Bibles: Nineteenth-Century Publishing and the British and Foreign Bible Society*, Cambridge Studies in Publishing and Printing History, Cambridge UP, 2002, pp. 1-34.

Hsy, Jonathan, and Rick Godden. "Universal Design and Its Discontents." *Disrupting the Digital Humanities*, edited by Dorothy Kim and Jesse Strommel, Punctum Books, 2018, pp. 91-112.

Kelemen, Erick, and Donald H. Reiman. *Textual Editing and Criticism: An Introduction*. Norton, 2008.

Library of Congress. "History." *National Library Service for the Blind and Physically Handicapped*. www.loc.gov/nls/about/organization/history/#two, accessed July 30, 2019.

McGann, Jerome. Interview by Thomas Hill. The Library Café. 8 Apr. 2015. library-cafe.blogspot.com/2015/04/jerome-mcgann.html, accessed May 8, 2015.

Museum of the Bible. "Pre-Braille Bible." www.museumofthebible.org/book/minutes/689, accessed July 30, 2019.

National Endowment for the Humanities. "Building an Accessible Future for the Humanities." 2013. www.neh.gov/divisions/odh/institutes/building-accessible-future-the-humanities, accessed April 4, 2014.

National Federation of the Blind. nfb.org.

Shep, Sydney J. "Bookends: Towards a Poetics of Material Form." *Teaching Bibliography, Textual Criticism and Book History*, edited by Ann R. Hawkins, Routledge, 2006, pp. 38-43.

Shillingsburg, Peter L. *From Gutenberg to Google: Electronic Representations of Literary Texts*. Cambridge UP, 2006.

Society for the History of Authorship, Reading, and Publishing (SHARP). sharpweb.org.

Text Encoding Initiative (TEI). tei-c.org.

United States Department of Justice. "The Americans with Disabilities Act." ada.gov.

Veispak, Anneli, Bart Boets and Pol Ghesquiere. "Differential Cognitive and Perceptual Correlates of Print Reading Versus Braille Reading." *Research in Developmental Disabilities*, vol. 34, 2013, pp. 372-385.

Watts, James. "How Books Matter: The Three Dimensions of Scriptures." *How and Why Books Matter - Essays on the Social Function of Iconic Texts*. Equinox eBooks Publishing, 2019, pp. 7-29.

Williams, George. and Erin Templeton. "Accessibility in Digital Environments." Digital Humanities Summer Institute, University of Victoria, Victoria, BC, Canada, Jun. 3-7, 2017.

Appendix: Generating Accessibly DH Output Using TEI-Encoded Uncontracted Braille

A Brief Braille Primer

Despite mass digitization, Braille remains relevant and useful to people who have print disabilities. As I mentioned earlier, Braille is a tactile writing system. It uses combinations of raised dots in groups of six, called a "Braille cell," to represent different characters, groups of characters, or words (Braille Institute). Figure 1 shows the configuration of the Braille cell and how the dots are numbered, and Figure 2 shows the basic Braille alphabet, which is called "uncontracted" or "Grade 1" Braille. This Braille, however, is very basic; different Braille alphabets and character sets exist for different languages, and even for mathematical notation. For example, the cell with dot 1 raised denotes the letter "a" in plain English, but the numeral "1" in mathematical notation. English Braille used to have separate character sets for American and British Commonwealth versions, but in 2016 these were combined into United English Braille (UEB).

Figure 1: The Braille cell with numbered dots

Figure 2: The basic, or Uncontracted, Braille Alphabet

Braille output is very bulky. It is generally possible to place only 40 or 50 cells on a line of 8 ½ x 11-inch paper or 12 x 12-inch paper, respectively. Thus, a single Braille book can take up several volumes, often assembled in large ring binders, and even when embossed on both sides of the paper (interpoint Braille). For example, a typical interpoint Braille Bible still occupies 20 binders or volumes and takes up over six linear feet of shelf space (Braille Bibles International). To mitigate the bulk, modern

Braille comes in two versions that are essentially longhand (uncontracted, formerly called “Grade 1,” as referenced earlier) or shorthand (contracted, formerly called “Grade 2”). In uncontracted Braille, a single cell with a specific raised-dot combination stands for each separate symbol in a character set (Figure 1). To produce a text in uncontracted Braille constitutes transliteration in that there is a character-for-character match between the original text and the transcription.

In contracted Braille, on the other hand, a single cell can denote an entire word. For example, in contracted Braille, the letter “w,” when embossed by itself, represents the word “with.” Contracted Braille also uses symbols to stand for parts of words, such as “-er,” “-ed,” and “-ing.” In some cases, dot configurations across two or even three cells represent words or phrases. Uncontracted Braille remains preferable in the present TEI-encoded text-production process because the TEI tags themselves must also be rendered in Braille to provide the reader with orientation to both the physical and semantic structure of the document. It is these tags that provide the descriptive information about a document’s layout and appearance that would otherwise be conveyed visually, such as the locations of stanza breaks. Conflicts arise, of course, when using two sets of abbreviated tagging systems (TEI and contracted Braille). For example, in contracted Braille, the letter “l” rendered by itself represents the word “like.” In TEI markup code, however, the letter “l” indicates a line break. Therefore, it is more clearly advantageous to spell out the word “like” and use the machine-readable “l” and “/l” to encode line breaks. The next section explains how to do that.

Making Accessible, Physical, Scholarly Editions Using TEI-Encoded Uncontracted Braille

In this section I detail the specific steps I created based on the software and hardware that I was able to acquire with my allotted funding. Product mentions do not constitute commercial endorsements; rather, they simply indicate those products I was able to purchase and use. As such, this process does come with a bit of a learning curve in that experienced Braille users would have to learn enough TEI tags to make sense of the output, whereas those familiar with TEI would, at the very least, have to learn the basic English Braille alphabet (Figure 1), a few other symbols, and a few notation differences between handwritten Braille and “computer Braille.” Some users would, of course, have to learn both. Granted, ostensibly, one could port a TEI-encoded file to audio, where the text-to-speech software would read aloud both the codes and the text. However, the voice synthesizer’s inability to reproduce certain phonemes and words

still remains, as in the cases of Old and Middle English. Finally, as I have argued earlier, porting to audio does not preserve embodied textual interaction or the cognitive and intellectual benefits thereof, including increased comprehension and retention of information (Barton, Sinclair and Agato 22).

Acquiring and Editing Archival TEI Files

Today, digital humanists can obtain TEI-encoded files of scholarly or literary works from peer-reviewed DH archives, many of them open access with Creative Commons licensing. In some cases, the unconverted TEI files, those that have not yet been formatted for web presentation, are linked and made available for direct download. Researchers can also request TEI-encoded files from project developers. In other cases, opening a document in a DH archive and then selecting “View Source” from the “Edit” menu in most browsers will reveal the encoded text, which can then be copied in whole or in part depending upon audience needs. Copyright considerations may apply in the latter case in that even though the work of literature encoded may be in the public domain, the project developer’s edition of that text, as manifested via their code, may be protected. Here, a request for permission to use is in order.

Because every TEI file is encoded differently, it may be advantageous to use a TEI and XML-capable text editor such as Oxygen or Notepad++ to remove code that the end reader does not need. A literary scholar, for example, may not need the complex header that begins each TEI document, while such would most likely be of use to a DH scholar. Likewise, the density and complexity of tags can also be edited, eliminating that markup not of interest to the end user. A book historian, for instance, might only be interested in those tags that describe the structure of the original physical document, while a mainstream historian might only be interested in those parts of the text that have tags for dates and geographic locations (geotagging). A student may only need basic tags, such as those indicating parts of the text (lines, paragraphs, and stanza breaks) or names of people and places. This ability to customize the TEI files enables end users who have print disabilities to focus more clearly on their areas of study within the given text. After editing, for successful Braille output, the edited TEI file should be saved as plain text. The following subsections of this paper detail my exact technical process for generating the accessible output.

Generating the Output: Refreshable Braille Display (RBD)

RBDs receive text input from a computer or other device such as a tablet or smartphone via Braille translation software, which converts the text to Braille and then sends it on to the RBD. A series of movable pins configured as Braille cells inside the RBD then move up or down to create the Braille characters. This functionality enables users to read the output line by line. Figure 3 shows the RBD hardware configuration for this project.

Figure 3: Hardware configuration for RBD.

Technical Specifications

Hardware

- Freedom Scientific Focus 40 Blue Refreshable Braille Display (40 cell display)
- Computer (desktop or laptop)

Software

- Robust Text Editor, such as Notepad++ or Oxygen
- Freedom Scientific JAWS Professional for Windows version 18.0 (screen reader acts as Braille translation software between the computer and the RBD).

Process

The following steps result in TEI-encoded uncontracted Braille output to an RBD using the technical specifications listed above:

1. Connect the RBD to the computer via its USB cable.
2. Power on both the computer and the RBD.
3. Open JAWS, if it does not automatically launch.
4. In the JAWS “Options” menu, make sure that Braille output is set to the correct RBD, in this case the “Focus,” as shown in Figure 4.
5. Click “Advanced” in the menu shown above to make sure that Braille translation is set to US English Grade 1, which is uncontracted Braille, as shown in Figure 5.
6. Open a text editor, then type or open a TEI file. Figure 6 shows use of Notepad++ with the TEI-encoded Shakespeare’s birth date and place.
7. Output analysis.

Figure 4: Launch JAWS Professional and Verify Under Braille Settings that the RBD is Set to the Focus, as per the Technical Specifications Above

Figure 5: Verify that JAWS is set to Translate to Uncontracted Braille

Figure 6: Open or Create a TEI Text File in a Text Editor

Figure 7 shows this project's RBD in more detail.

Figure 7: TEI-encoded Uncontracted Braille Output to the RBD

Figure 7 shows the beginning of the output of the TEI text file on the RBD. Attention to detail here reveals that the RBD's Braille cells have eight dots, not six. There are two reasons for this configuration. First, some new UEB symbols use eight dots. But when the translation software, in this case JAWS Professional, is set to output uncontracted or Grade 1 Braille, the extra two dots at the bottom of each cell house indicators. As mentioned earlier, uncontracted Braille rendered by computer can sometimes render the Braille code differently.

In Figure 7, for instance, the first character displayed is the left angle bracket. The two pins raised below the angle bracket symbol indicate the position of the cursor. These pins move up and down repetitively to help the user keep their place. Following the next left angle bracket, note the next character. It is a capital "P" beginning the TEI code "PersonName". Here, on the RBD, we do see an indicator pin raised to mark the capital letter. In handwritten Braille, this dot indicating the capital letter would appear in the previous cell with dot six in the raised position. This type of output offers users many benefits, as described in the next section.

Benefits

TEI-encoded uncontracted Braille files ported to an RBD offer scholars who have print disabilities many advantages, including the following:

1. The portability of TEI text files as simple, .txt email attachments means that accessible DH materials can be very quickly put into the hands of users who need them, provided they have the necessary working knowledge of both codes, Braille translation software and access to an RBD.
2. Electronic files mean fewer bulky printouts.
3. At least in the case of the present project, JAWS and the Focus 40 RBD offer better rendering of both TEI and Braille codes in terms of indicators for capitalization and punctuation.
4. RBDs are very portable; scholars who have them, and the necessary translation software on their laptops, can easily carry their access with them as they attend conferences or visit archives or libraries.

Challenges

RBD users can also face some challenges, as follows:

1. There is, of course, the initial learning curve. A scholar new to Braille will need to learn the basics, and the same applies to scholars new to TEI.
2. RBDs are expensive, but the cost is steadily coming down due to advances in technology.
3. RBSs are very fragile with many moving parts. Repair can be costly or even cost-prohibitive and can also take a very long time.

On balance, however, an RBD coupled with a good text editor and uncontracted Braille translation software can certainly help to remove visual access barriers some scholars face. The compact portability of both the TEI files and the Braille equipment further facilitate this access.

Generating the Output: Braille Printer

Similar to RBD output, TEI-encoded uncontracted Braille files can be ported to a Braille printer. The Braille translation process via bridging software works in the same way as for the RBD output process described above.

Braille printers are available in two output methods. Traditional “striker” or “embosser” type Braille printers work much like the RBD, except that the raised pins more forcefully emboss the raised Braille dots into special, thick paper, which is the consistency of cardstock. In fact, regular craft store cardstock works fine in these printers in the absence of Braille paper. The primary advantages of this type of printer are speed and the ability to produce high-volume output. The drawbacks are also

similar to the RBD in that these printers are costly, fragile, and difficult to repair. Furthermore, they are very loud, the sound having been likened to that of a Gatling gun or automatic rifle. Some Braille embossers even come with noise-reduction covers.

Thermoform Braille printers, on the other hand, work more like modern 3D printers. The “ink” is a plastic substance similar in consistency to hot glue. These machines print the Braille onto special paper that is also thick and somewhat sticky. These printers are cheaper and quieter; however, only the special thermoform Braille paper can be used. Plus, there is the extra cost of the thermoform “ink.” The disadvantages of thermoform printers are similar to those of old-school inkjet printers. The “ink” is expensive and creates an unpleasant odor when combined with the paper. Printouts are very hot to the touch and must be laid flat to dry. The hot-glue-type Braille dots are flatter and thus harder to read, and they tend to flake off with use and age.

Either type of Braille printer, embosser or thermoform, can be augmented to also print regular text in regular ink along with the Braille. Some Braille printers of both varieties require paper to be fed through twice, once for regular ink and then once for the superimposed Braille. Other printers print one line of regular text followed by the same line in Braille. The latter configuration facilitates document editing or grading. This is by far the most versatile type of Braille printer, especially when sighted people who are unfamiliar with Braille, such as professors, archivists, and librarians, need to carry out the production of Braille documents because they can at the very least check the TEI code by eye.

Technical Specifications

Hardware

- ATC (American Thermoform Company) Braille & Print printer (refurbished, purchased from Freedom Scientific, interlinear, prints lines of regular text in regular ink with unoccluded Braille below)
- Computer (desktop or laptop)

Software

- Robust Text Editor, such as Notepad++ or Oxygen
- IndexBraille text-to-Braille translator

Process

The following steps result in TEI-encoded uncontracted Braille output from a Braille printer using the technical specifications listed above. These steps assume that the printer is connected to the host computer with drivers installed and that it has enough ink and paper for the desired output.

1. Open or create a TEI file in a text editor. Save the file as .txt.
2. Close the text editor.
3. Open the Index Braille translation and printer utility on the host computer's desktop.
4. In the "Options" menu, accessed in the lower left corner of the dialog box, make sure the translator is set to en_us_g1 for uncontracted Braille.
5. Click the "Add" button in the upper right corner of the utility dialog box.
6. Navigate to the .txt file to be printed and select it. The filename should appear in the main window of the dialog box, and will be highlighted in blue.
7. Click "Emboss" in the lower right corner of the print utility dialog box.
8. Retrieve printout from the ATC Braille & Print.

Figure 8 shows the ATC Braille & Print embosser used for this project and its control panel in detail. The benefits of its affordances appear in forthcoming discussion. Figure 9 shows a bit of the printed output of a TEI encoding template.

Figure 8: Functions of the ATC Braille & Print

Figure 9: TEI-encoded Uncontracted Braille Output to ATC Braille & Print Embosser

Benefits

Two primary benefits ensue from the availability of a physical TEI-encoded document printed in uncontracted Braille:

1. The physical printout affords the reader more ability to explore a document by going back and re-reading if necessary. While this is possible with the RBD, that device, unlike the printout, does not allow, for example, for the placement of tactile markers such as tape flags in the document to mark a reader's place.
2. In the interest of the material text, the primary advantage of a Braille printout of a scholarly work is that it preserves the physical interaction with text.

Challenges

Despite the very significant benefits noted above, challenges remain in producing embossed TEI-encoded uncontracted Braille. These disadvantages include the following:

1. As noted earlier, Braille embossers are costly, noisy and difficult to repair.
2. Braille output can be inaccurate, as in the case with the RBD, in the absence of a robust literary Braille translator.
3. Braille documents remain bulky, and must be mailed to scholars, or the scholars must travel to a location with available equipment to use it.

On balance, though, these processes provide access to scholarly digital materials that would otherwise not be available.

Discussion

This section takes up some additional points of discussion not covered in the technical analyses above. Granted, both TEI and Braille present learning curves. Other project dependencies include the cost of equipment and software and even the very use of Braille itself. On the latter point, an uninformed argument posits that Braille is dead, killed off by so many talking devices that blind people don't bother to learn or use it anymore. According to the National Federation of the Blind, one of three leading organizations of people who have significant vision impairments in the United States, only one in ten know how to read Braille. However, although this organization does define blindness as having a visual acuity of 20/200 or less, or a visual field of less than 20 degrees, in the better corrected eye, it does not define what it means by "know Braille" or "read Braille." My own lived experience, now 50 years legally blind, and my lifelong interactions with other people who have low or no vision, contradicts the statistics.

Braille is very much still in use; a walk down the hall to any public elevator in the United States will bear this claim out. By this I mean that while many people who have visual access needs do turn to other technologies for long reading, most of us do in fact still use the basic, uncontracted Braille on a daily basis for labelling things, making quick notes, and, yes, detailed reading when the details matter. This is especially true of academic professionals with low vision. For example, while an audio recording of the Great Books trade edition of *The Canterbury Tales* might be fine for a high school student who needs to know only the main characters and who said what to whom, this non-scholarly edition will certainly not do for the graduate student earning a Ph.D. in medieval literature. Even audiobooks of scholarly editions too often omit the apparatus, such as notes and annotations, that make the edition scholarly in the first place.

The present research into accessibility in digital environments reveals as a collateral benefit that the DH community stands in a prime position presently to offer better texts when scholars need them. Even using a screen reader to access the display version (non-TEI version) of a text in a DH archive provides a more rigorous scholarly experience of that text than a trade, all too often abridged, audiobook. But as noted above, when the material text matters to the scholar, TEI-encoded uncontracted Braille delivers. For example, in addition to gathering notes at the end of an audio text,

making accurate referencing nearly impossible, other textual elements such as glosses or stage directions often get left out as well, or, worse, awkwardly read into lines of text such that the reader's flow and understanding suffers constant interruption. The low-cost alternative of simply porting a TEI-encoded text file to a screen reader and listening to it also proves ruefully unworkable in that screen readers notoriously mispronounce even modern English words, let alone Middle English, and spelling can only be accessed by setting the screen reader to read one character at a time, again destroying reading flow and comprehension.

Another challenge to the premises of this work comes from the advent of UEB in that it may supplant both traditional contracted and uncontracted Braille. Again, the argument for replacement is uninformed; my own experience evaluating UEB on several platforms is that it actually includes functionality for uncontracted Braille by setting any future devices to default to it to "Grade 1 mode." UEB is new, having come online in 2016, and so most Braille devices such as printers and RBDs still do actually default to contracted Braille. This is a known challenge, but easy to fix by simply changing the settings. There remains, however, the need to inform Braille product developers not to discontinue support for uncontracted Braille.

Footnotes

1. Mentions of products and services in this article do not constitute commercial endorsements. [↵](#)
2. Technical assistance for the process described in this article is available from the Center of Digital Humanities Research (CoDHR) at Texas A&M University. For more information, please visit <https://www.codhr.tamu.edu>. [↵](#)