10.

Digital editing and publishing in the twenty-first century as a cooperative for small-scale editions

Juniper Johnson, Serenity Sutherland, Neal Millikan and Ondine Le Blanc

Introduction

In the past decade, documentary editors and the organisations that fund them have worked to broaden the digital publishing options available to small digital editions and to diversify the types of projects that receive funding. One key intervention towards this goal is the digital publishing platforms initiative funded by the National Historical Publications and Records Commission (NHPRC) and the Andrew W. Mellon Foundation. In 2018, after a round of planning grants, the NHPRC and Mellon funded three different cooperatives, each with the intention of producing pathways to publication for small to midsize digital editing projects. Thinking of the results of these publishing platforms as a 'sustainable digital edition publishing ecosystem', the working group that imagined the grants' structure sought publishing platforms that offer greater interoperability, sharing across institutions and project editors, and increasing access to records that are free, usable and able to facilitate new research and learning.[1] As members of one of these three digital publishing

1 R. Darrell Meadows, 'Building A Sustainable Digital Edition Ecosystem',Scholarly Communication Institute, 18 May 2016.

cooperatives, the authors argue that this type of publishing model is transformative and points to the future of digital editing in the twenty-first century. More than simply meeting the requirements outlined in the grant, this type of cooperative publishing will be an essential part of documentary editing praxis in the future as it allows for a diversity of voices and editorial approaches that the field of scholarly editing greatly needs.[2]

In what follows, we describe what cooperative publishing is and how it is transformational to the making of an edition. The power of cooperative publishing is three-fold: (1) the sharing of resources, both financial and structural; (2) the collaboration of content expertise across a wide range of topics; and (3) the support of a community striving for the same goal.[3] With this opportunity to create a scalable digital publishing platform, the cooperative participants engage in the process of re-imagining the digital edition and digital scholarly workflow. Our work proposes a publication model designed around a deliberate praxis: one that is collaborative, equitable and designed from the bottom up.[4] The cooperative aims to reduce the barriers to publishing for all, especially for editors lacking institutional support.

2 The scholarship on decolonising the digital humanities and the archive inform this perspective on the power of a cooperative publishing platform. See Roopika Risam, *New Digital Worlds: Postcolonial Digital Humanities in Theory and Praxis* (Chicago: Northwestern University Press, 2018); and Christina Boyles, Andy Boyles Petersen, Elisa Landaverde, and Robin Dean, 'Postcustodial Praxis: Building Shared Context through Decolonial Archiving', *Scholarly Editing* 39 (2021).

3 Scholars have discussed what collaborative digital publishing might look like, and what we propose here builds on the principles of interdisciplinary collaboration through the mechanism of a cooperative. See Peter Robinson, 'Some Principles for Making Collaborative Scholarly Editions in Digital Form'. *Digital Humanities Quarterly* 11, no. 2 (May 2017).

4 Kathryn Simpson and Heather F. Ball, 'Editing to Avoid Exclusion: Understanding the Subjective Power Dichotomies in Scholarly Editing', *Scholarly Editing* 39 (2021).

Background

The purpose of the Primary Source Cooperative (PSC) at the Massachusetts Historical Society (MHS) is to provide a platform, designed and governed by consensus, to assist with the digital publication of documentary editions led by scholars who study the American long nineteenth century (1789–1914) and who would not otherwise have a portal for online publishing that is affordable and supportive. The work of the PSC will benefit digital publishing more broadly and the public generally, since documentary editions have, historically, performed a vital cultural role by translating primary source materials into formats that users can find, read and understand. With this focus, the PSC is a resource for scholars and other users seeking primary sources about this critical period in American history, when revolution and reform were causing fundamental changes in social and political culture. More broadly, however, the PSC's praxis is meant to be reproducible, making available a model of digital publication that runs on human and organisational collaboration that can be adapted to varying circumstances and replicated for use by other cultural institutions and archival repositories. Working towards the goal of a federated network of cooperatives, we see our cohort as one in a landscape of overlapping systems, each with its own topical parameters, administrative arrangements, tools, infrastructure and financial models. The existence of a plethora of editor-driven cooperatives is a crucial step towards realising the rich and adaptive environment needed to improve the generation of new editions and the growth and accessibility of documentary editing.

The impetus for the cooperative model for digital publishing, and the anticipated network of cooperatives, comes from a pressing need to distribute representations of archival materials. Two challenges present themselves: first, the issue of institutional support (both structural and financial); and second, a readily available pathway to publication and dissemination. In regard to the first challenge, the editors at each of the four editions represent a wide variety of institutional and academic settings, but overwhelmingly the editors lack strong institutional support to sustain digital editing

work. Some of the editions have been successful in earning grants from the National Endowment for the Humanities; however, as digital humanities scholars have noted, researchers endeavouring to complete such projects at small to midsize institutions face challenges relating to resources and funding, especially if the institution lacks a dedicated Digital Humanities space.[5] The PSC thus serves as a potential home for digital editions wishing to employ digital humanities tools but lacking the institutional support or a dedicated DH-lab at their own institutions to do so.

The PSC currently contains four editing projects: the John Quincy Adams (JQA) Digital Diary, the Papers of Roger Brooke Taney, the Catharine Maria Sedgwick Online Letters (CMSOL) and the Ellen Swallow Richards Papers. One of America's great statesmen, John Quincy Adams's (1767–1848) distinguished career in public service spanned six decades and included roles as diplomat, secretary of state, president and congressman. The John Quincy Adams Digital Diary makes JQA's diary, which spans over 68 years, truly accessible for the first time. The Papers of Roger Brooke Taney (1777–1864), a project based at the University of West Florida in Pensacola, will digitally publish annotated transcriptions of Taney's papers (correspondence, legal documents and so on). Each online volume will capture a separate aspect of Taney's life and career, including his tenure as chief justice of the United States Supreme Court (1836–64) and his family life. During her lifetime, Catharine Maria Sedgwick (1789–1867) became known in the United States as the most significant, experimental, influential and highly regarded woman writer in the Early National period of American literature. The Catharine Maria Sedgwick Online Letters (CMSOL) project makes freely available authoritative transcriptions of all surviving letters written by Sedgwick

5 Bryan Alexander and Rebecca Frost Davis, 'Should Liberal Arts Campuses Do Digital Humanities? Process and Products in the Small College World', *Debates in the Digital Humanities*, ed. Matthew K. Gold, (Minneapolis: University of Minnesota Press, 2012), 368–89; Peter Robinson, 'Digital Humanities: Is Bigger Better?', *Advancing Digital Humanities*, ed. Paul Longley Arthur and Katherine Bode (London: Palgrave Macmillan, 2014), 243–57.

during her nearly seven decades as an active correspondent. At the end of the nineteenth century and in the early part of the twentieth century, Ellen Swallow Richards (1842–1911) worked to broaden women's access to education and science. She was the first woman to graduate from and then teach at Massachusetts Institute of Technology, a contributor to the twentieth-century home economics movement and a chemist.

Our cooperative is supported by two institutional centers. First, the MHS provides the technical aids that make it possible for editors to prepare content (transcriptions and contextual material) as predictable XML, the PSC's baseline for its digital publication system. Second, the Digital Scholarship Group (DSG) at Northeastern University (NEU) transforms source data into digital derivatives (such as visualisations and contextual data) generated from the editions that co-exist within a much larger pool of data gathered from other archival collections. Technical specialists at NEU have built data tools that feature network visualisations of individuals mentioned within the letters, prevalence of subject headings and sentiment analysis of the data. Working together on design and access, the MHS team and the editors are building the PSC website, including the content management system that enables editors to upload and manage their own editorial content. End users, the general public and scholars will access, read and search the content of the digital editions through direct engagement with the transcribed documents, visualisations and a database of annotations generated from each edition's data.

Assisted by the institutional support of the MHS and NEU, the editions not only publish documents, but also strive to produce data that is usable by scholars as well as visualisations based on that data that are accessible by the public. Each edition participates in the building process, including the review of publishing tools, writing of governing documents, and aesthetics of website design. In this way, the editors themselves have direct say over not only the content of their edition, but the design and functionality as well. The cooperative worked with several consultants on various technical and

content-specific aspects. At every stage of the design process, cooperative members have had an opportunity to give feedback and suggest changes to website display, data visualisations, design of individual pages, database architecture and usability, and overall site appearance and functionality. Each editorial team has the ability to manage how their individual site will display.

Many editors are subject-matter experts who have years of experience researching and writing on the historical figures whose papers they are editing. They are also highly adept at handling and interpreting archival materials, but may not have access to the tools and infrastructure needed to publish their work online – and especially not in a form that makes the best use of their insight as scholars or their understanding of the source materials as editors. The impediments these editors face are thus also a detriment to the researchers who would benefit from the content of these source documents, since the digital medium is the first – and sometimes only – point of access people use to find historical sources. Working together to create an alternative pathway, the PSC is devising tools and organisational structures that aim to maintain editorial agency in the semantic markup of documents, based on the principle that editors must have a defining role in preparing electronic text.

Collaboration as praxis

The collaboration and sharing of expertise is at the heart of the PSC's mission. Since 2017, our goal has been to pool our resources – financial, technical know-how, documentary editing experience, and subject matter knowledge – to create a viable and sustainable digital publishing platform. Each project in the cooperative combines human power and resources to create an online portal that is strengthened by representing multiple voices and that creates a supportive environment in which editors can work and receive feedback when they have questions/issues. The theoretical and practical advantages and challenges of collaborative, interdisciplinary digital projects have constituted a central debate and driving conversation

in digital humanities and digital scholarship more broadly.[6] Despite what Claire Warwick outlines as the 'long-running debate about whether the creation of a digital resource is "just" a service task or whether it has an essential intellectual component',[7] we wish to contextualise collaboration as interdisciplinary praxis, paying attention to the impact of building in collaboration from the beginning of a project throughout its life, working to create structures that will yield sustainable, replicable digital publishing opportunities for scholars at institutions without existing infrastructure. Whether this means 'translating' disciplinary language and methods into forms accessible to others in interdisciplinary projects or working in tandem to create new ways of communicating and collaborating, such scholarship is tremendously valuable.

6 Since the emergence of digital humanities as the 'next big thing' (as described by William Pannapacker in his blog 'The MLA and the Digital Humanities' in The Chronicle of Higher Education) in the late 2000s to early 2010s, a central genre of scholarship to the field has been works-in-progress. Discussing the practical, routine work of how digital projects get done, scholars have keenly explored how disciplinary boundaries are overcome (or adapted to) in such interdisciplinary work. As digital scholarship has become more mainstream across humanities disciplines, these discussions have grown into reflections on project design and implementation, fuelled by scholars increasingly paying attention to frameworks from feminist and queer theory, disability studies, critical race theory, and so on, in their digital work. This progression of digital humanities is broadly discussed in the three editions of *Debates in Digital Humanities* (University of Minnesota Press, 2012; 2016; 2019). For discussions specifically about how disciplinary content expertise adds or challenges such endeavours, see Lisa Spiro, '"This is Why We Fight": Defining the Values of the Digital Humanities', *Debates in Digital Humanities*, ed. Matthew K. Gold (University of Minnesota Press, 2012) and Julia Thompson Klein, *Interdisciplining Digital Humanities: Boundary Work in an Emerging Field* (Ann Arbor: University of Michigan Press, 2015).

7 Claire Warwick, '"They Also Serve": What DH Might Learn about Controversy and Service from Disciplinary Analogies' *Debates in the Digital Humanities*, ed. Matthew K. Gold and Lauren F. Klein (Minneapolis: University of Minnesota Press, 2019), part 1, chapt. 4, https://dhdebates.gc.cuny.edu/projects/debates-in-the-digital-humanities-2019.

A central component of scholarly discussion on digital humanities work and collaboration is *how* we talk about, categorise and understand such tasks: what words, metaphors or ideas are we using to describe the intellectual and physical labour of digital work? Whose work is being discussed? Such analyses, like Julia Flanders's focus on 'building' in the ties between individualistic maker culture as opposed to building 'otherwise' with feminist, collaborative efforts,[8] Bobby L. Smiley's exploration in job titles, disciplinarity, and labour of DH librarians,[9] or Jacqueline Wernimont's dissection of objectivity in digital methods,[10] emphasise the importance and need for critical attention to how we understand and share the work of digital projects across fields and other institutional boundaries. In the cooperative, this sharing across disciplinary boundaries is well represented as editors come from backgrounds of history, literature, political science and communication studies.

Our cooperative's interdisciplinary collaboration and sharing of resources occurs both within and beyond the PSC. While multiple barriers exist to creating a digital edition (financial and technical barriers being the most insurmountable for many academics), we believe that the digital publishing model developed by the PSC can be replicated by other cultural institutions. With this idea in mind, the work we have done and the decisions we have made in order to create, sustain and grow our cooperative has maintained an eye towards usability by other projects in the future. Documenting our actions and the tools we create have been essential so that future editors may

8 Julia Flanders, '"Building Otherwise", *Bodies of Information: Intersectional Feminism and Digital Humanities'* ed. Elizabeth Losh and Jacqueline Wernimont (Minneapolis: University of Minnesota Press, 2018), 298.

9 Bobby L. Smiley, 'From Humanities to Scholarship: Librarians, Labor, and the Digital', *Debates in the Digital Humanities*, ed. Matthew K. Gold and Lauren F. Klein (Minneapolis: University of Minnesota Press, 2019), part IV, chap. 35, https://dhdebates.gc.cuny.edu/projects/debates-in-the-digital-humanities-2019.

10 Jacqueline Wernimont, 'Introduction: Methods for this History of Quantum Media', *Numbered Lives: Life and Death in Quantum Media* (Cambridge: MIT University Press, 2020), https://covid-19.mitpress.mit.edu/pub/v3qjp2k8.

benefit from the workflows and processes undertaken by the PSC. Much of the collaborative work PSC members undertake is translatable to other editing projects. For example, one area in which our projects' editors have shared – and continue to share – our knowledge across the cooperative (and one that would also be useful to editors at large) is with our lists of standardised terms (for people and historical topics) created to assist in centralised annotation. From the beginning of the PSC, each member project agreed to identify and encode the individuals mentioned in their documents, and to encode all relevant historical topics that apply to a text. Creating both the lists of people and historical topics has truly been a group effort, requiring us to work together collectively among the editors as well as with external consultants to craft a standard, usable taxonomy for America in the long nineteenth century with all of its varied changes and challenges.

The final web development step in the Names effort was the creation of a shared database that would meld all of the individual entries from every edition-specific spreadsheet. The process of importing the data from four spreadsheets into one database required collaboration on the overlapping entries, in order to avoid conflicts, as well as confirmation that names that appear to be 'duplicated' in more than one spreadsheet are actually the same person. (Anglo-Americans of the time had a tendency – irksome for historians – to re-use the same names frequently.) Finally, the team at the MHS and the editors coordinated the timing of the 'ingests', when the spreadsheet information made its official passage into the centralised cooperative-wide database.

Once that stage was completed and the spreadsheets became irrelevant, editors shifted their work to the dashboard that will become their primary work environment for managing the names data; uploading, reviewing and publishing their XML content; and – still in the future – managing the historical topics they have tagged in each of their documents. This dashboard is Beck's customisation of a standard WordPress interface; WordPress provides the foundation for the web content management system. In the Names UI, editors can search for and edit existing names records or create new records

when necessary. Beck's system automatically assigns each new name a unique HUSC (Hyphenated Unique String of Characters) based on the family and given name information keyed in by the editor. This important feature prevents the repetition of a HUSC within the database. The names database also facilitates the encoding of personal names with the inclusion of a mechanism by which an existing HUSC can be copied and pasted in an XML file simply by clicking anywhere on the HUSC and then pasting the HUSC inline where that individual is mentioned. This is a vast improvement over the encoding workflow when working from the names in spreadsheets where the full HUSC had to be highlighted, copied and pasted, and occasionally led to encoder error within an XML file.

One challenge with this work was that projects often first came across individuals at different stages of their lives; for example, a young woman mentioned in JQA's diary in the 1790s might be the same woman who went by a married name in an 1820s Sedgwick letter. The intellectual work of determining which HUSCs did or did not overlap provided editors with an opportunity to revisit and refine an individual's record, adding other pertinent information to a record to disambiguate or, in the opposite case, combine individual projects' HUSCs. The result of this work is that the database containing individuals mentioned in the documents now contains the unique HUSCs from all four member editions as well as any new name records that are being created. It also provides a platform into which every editor can add to and refine the contents of a name record. Cross-edition searching would not be possible if each of our digital editions were siloed on individual websites rather than being part of one unified web portal, pooling the research and intellectual work of multiple editors to help improve a record, both for use by the cooperative editors and for our website users.

Although it is still in an earlier stage of development, the same process exists for our lists of historical topics: each project started with their own list, then we began collectively reviewing our lists, providing explanatory text on when and how we utilised a topic during analysis and determining what topics related to multiple projects and where we could

adjust our terminology to potentially combine topics. While some of the topics are unique to a given edition, such as Family Residences (Adams Family), other topics overlap the editions (such as Immigration and Science and Technology). The intentionality we put into crafting our lists of historical topics lends itself to a fuller type of search capability for users beyond the traditional keyword search. The creation of a PSC topics list depends on the direct participation of the editors, who have the requisite subject expertise to recognise references to specific topics even when these standard-ised terms are not specifically elucidated in the historical text. This potential for cross-edition searching for historical topics by users is a significant feature of the cooperative.

While our cooperative is new, we are already seeing the potential for interoperability. For example, members of the PSC, in discussion with representatives from the Center for Digital Editing at the University of Virginia, discussed how shared taxonomies would greatly assist both cooperatives' future editors and website users. Future editors, no matter the cooperative they belong to, could craft a list of relevant topics for their own projects from a shared taxonomy without having to take the time we did to create a list of subjects from scratch. Utilising controlled vocabularies across editions would also be a boon for researchers, who could search for and find the same term being utilised to represent a common topic in multiple digital editions. Another strategy we are exploring at the PSC is the idea of utilising cross-references within our topics lists to redirect website users to terms that they may be looking for in a document, but which we have chosen to represent with another word or phrasing; an example is 'westward exploration', 'westward expansion', and 'westward migration'. Once a decision is made on the term to be used, other instances of the phrase could still direct users to documents encoded with the term using 'see' and 'see also' search results.

Beyond creating databases of historical topics and personal names, PSC edition partners, in concert with MHS staff, develop and maintain editorial standards for their own projects. The cooperative has a larger set of standards that each project must adhere to in order to maintain baseline standards of editorial quality; however, some

flexibility is allowed. For example, each project develops an editorial statement that outlines unique practices to each edition, such as arbitrary devices used for clarifying texts, transcription and verification policies and other unique facets.

Some collaborative practices of the PSC are beneficial solely to the group of editors gathered at the moment. For example, we regularly share resources during monthly Zoom 'editorial hours'. These meetings are an opportunity for all of the editors to come together and discuss any questions or issues that have arisen in their own editions or that relate to the larger cooperative community. These editorial hours provide a space where members use fellow editors as a 'sounding board' and grapple with editorial or markup questions that we are unsure how to handle ourselves. Editorial work can often be isolating, especially for lone editors and/or editors working on a small project, and the editorial hours offer community and avenues for collaboration.

Cooperative members also determine the PSC's governance structure: drafting a constitution, by-laws, and mission, vision and values statements; establishing a Governance Board; formulating the review process for accepting new member editions; and setting down the peer review system for assessing the quality of edition content prior to full publication. The governance documents structure the PSC so that every project has equal weight, both in terms of making decisions and in terms of doing the executive and administrative work necessary to keep the cooperative moving forward. Cooperative members also proposed potential business models to provide financial stability after grant funding ends. The business plan is as yet a flexible document, but one strategy for member inclusion is having each project pay a modest fee (based on edition size/grant awards and so on) as a subvention – analogous to institutional subventions for page fees – in order to maintain membership. Structural components such as robust governing documentation will help sustain the momentum of the PSC. Other institutions might follow our model and host digital editing cooperatives in the future; thus, we need to ensure the viability of our own cooperative to show this model is replicable by other organisations. Governance documents will be

published on the PSC website and will be available to the documentary editing community.

Critical application: Word Enhancement Template (WET) to XML workflow

The PSC's infrastructure and workflows are built to address the ongoing challenge of balancing the technical needs of creating digital editions (for example, content production with transcription, processing and versioning), content expertise and accessibility for all involved parties, regardless of background technical skills. Like any collaboration of this scale, there is a wide range of skills and expertise brought to the table. As previously outlined, the PSC has three distinct working groups: four teams of documentary editors, the publications and web development team at the MHS and the database and digital derivatives specialists at the DSG. Individually each group has their own workflows that, when connected, form a complex ecosystem of data creation, curation, collation and computation.

Since the beginning of the PSC in 2018, we have focused on developing and documenting editorial policies that balance each edition's individual priorities with shared cooperative-wide technical requirements for predictable content processing and pipelines. Additionally, we are developing tools for content creation that reflect the working preferences and intellectual value of the editions: building, testing, and refining practices as the editors use and reflect on them. A central goal of this project is to address barriers editors face, and working with XML can be a large barrier. The codelike nature of markup languages and XML editing software programs are intimidating and inaccessible to some. In response to this ongoing challenge and significant editorial need, the PSC has created a tool uniquely adapted to its immediate users through their direct and ongoing participation throughout its development. This workflow (aka WETVAC) begins with a structured transcription template in Microsoft Word and transforms it with a script (using XSLT and JavaScript) into well-formed, consistent XML.

The PSC transcription template (called the Word Enhancement Template or, more broadly, WET documents) utilises the template feature of Microsoft Word to allow editors and any assistants, graduate researchers or undergraduate students involved with the project to transcribe documents in a familiar word processing environment while also 'marking' document structures through pseudo-markup elements we call *markers*. In WET files markers are set apart from the other text of the document with beginning and ending double curly brackets (acting similarly to the starting and closing tags of XML elements) to identify and distinguish key documentary structures or metadata fields (see Figure 10.1). Thus, WET markers are used to 'encode' metadata fields (author and editor names, transcription dates, subject headings) and document structure, particularly reflecting the most common semantic traits of the nineteenth-century manuscripts, such as date-lines, salutations, paragraphs, datelines, postscripts and so on. Additionally, WET utilises Word's default formatting to 'encode' textual features like superscript characters and strikethroughs, where an author indicated the desire to delete text. WET was also created to allow editors to comment and annotate documents during the transcription process through markers like {{COMMENT}} and {{NOTE}} (see Figure 10.1). All the formatting and markers present in a WET document are there to assist the transformation to corresponding XML elements in the next stage of the document transformation workflow.

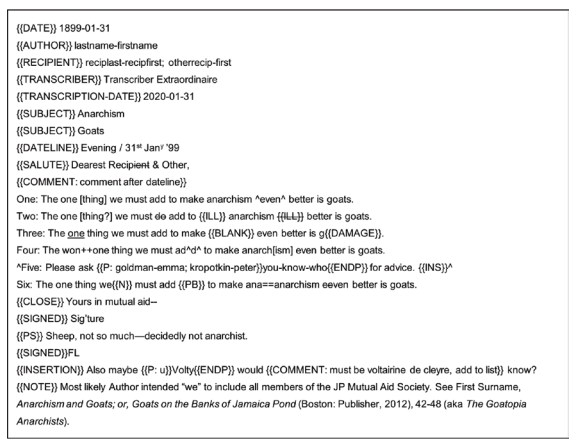
{{DATE}} 1899-01-31
{{AUTHOR}} lastname-firstname
{{RECIPIENT}} reciplast-recipfirst; otherrecip-first
{{TRANSCRIBER}} Transcriber Extraordinaire
{{TRANSCRIPTION-DATE}} 2020-01-31
{{SUBJECT}} Anarchism
{{SUBJECT}} Goats
{{DATELINE}} Evening / 31st Janʸ '99
{{SALUTE}} Dearest Recipient & Other,
{{COMMENT: comment after dateline}}
One: The one [thing] we must add to make anarchism ^even^ better is goats.
Two: The one [thing?] we must de add to {{ILL}} anarchism {{ILL}} better is goats.
Three: The one thing we must add to make {{BLANK}} even better is g{{DAMAGE}}.
Four: The won++one thing we must ad^d^ to make anarch[ism] even better is goats.
^Five: Please ask {{P: goldman-emma; kropotkin-peter}}you-know-who{{ENDP}} for advice. {{INS}}^
Six: The one thing we{{N}} must add {{PB}} to make ana==anarchism eeven better is goats.
{{CLOSE}} Yours in mutual aid--
{{SIGNED}} Sig'ture
{{PS}} Sheep, not so much—decidedly not anarchist.
{{SIGNED}}FL
{{INSERTION}} Also maybe {{P: u}}Volty{{ENDP}} would {{COMMENT: must be voltairine de cleyre, add to list}} know?
{{NOTE}} Most likely Author intended "we" to include all members of the JP Mutual Aid Society. See First Surname, *Anarchism and Goats; or, Goats on the Banks of Jamaica Pond* (Boston: Publisher, 2012), 42-48 (aka *The Goatopia Anarchists*).

Figure 10.1 Excerpt from a fictional WET document created to test the WETVAC output for consistency. Source: Authors.

When completed, each WET file is fed into WETVAC, an online script with a drag and drop user interface developed by MHS web developer Beck.[11] Colloquially named after a vacuum cleaner, WETVAC annihilates the excessive encoding that underlies every .docx file, retaining only the human readable text and those metadata and structural features that are 'tagged' in the WET document by editors. As it extracts the text and marked structures, WETVAC converts the Word file into an XML file using a customised schema following the Text Encoding Initiative (TEI) guidelines (see Figure 10.2). Like other projects that utilise TEI as the

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-modelhref="https://www.primarysourcecoop.org/publications/pub/schema/primarysourcecoop_rev2.2.rng"
type="application/xml"schematypens="http://relaxng.org/ns/structure/1.0"  ?>
<?xml-stylesheettype="text/css"href="https://www.primarysourcecoop.org/publications/pub/css/authorview_rev2.css"   ?>
<TEI xmlns="http://www.tei-c.org/ns/1.0" xml:id="GOA00001">
  <teiHeader>                                                                                                    [88 lines]
  <text>
  <body>
  <div type="docbody">
    <opener>
      <dateline>Evening / 31<hi rend="superscript">st</hi> Jan<hi rend="superscript">y</hi> '99</dateline>
      <salute>Dearest Recip<del>ient</del> &amp; Other,</salute>
    </opener>
    <p>One: The one  <unclear>thing</unclear> we must add to make anarchism  <add>even</add> better is goats. </p>
    <p>Two: The one  <unclear cert="low">thing</unclear> we must  <del>do</del> add to  <unclear/> anarchism
<del><unclear/></del> better is goats. </p>
    <p>Three: The <hi rend="underline">one</hi> thing we must add to make  <space/> even better is g <gap/>.</p>
    <p>Four: The <choice><orig>won</orig><reg>one</reg></choice> thing we must ad<add>d</add> to make
anarch<unclear>ism</unclear> even better is goats.  </p>
    <p><add>Five: Please ask  <persRef ref="goldman-emma;kropotkin-peter" >you-know-who</persRef> for advice.  <ptr
type="insRef" n="1" target="GOA00001-ins1"/></add></p>
    <p>Six: The one thing we <ptr type="fn" n="1" target="GOA00001-fn1"/> must add  <pb n="2"/> to make
<choice><abbr>ana</abbr><expan>anarchism</expan></choice> e<del>e</del>ven better is goats. </p>
    <closer>
      <salute>Yours in mutual aid-- </salute>
      <signed>Sig'ture</signed></closer>
    <postscript>
      <p>Sheep, not so much—decidedly not anarchist.  </p>
      <signed>FL</signed>
    </postscript>
  </div>
<div type="insertion" xml:id="GOA00001-ins1">
  <p>Also maybe  <persRef ref="u">Volty</persRef> would <!--COMMENT: must be voltairine de cleyre, add to list  --> know? </p>

</div>
```

Figure 10.2 Sample of the output from a WET document converted to XML through WETVAC. Source: Authors.

standard markup language for representing textual data, the PSC built on the general TEI guidelines and developed specific encoding work-

11 WETVAC is hosted online on the PSC website at http://primarysourcecoop.org/tools/wetvac/. The webpage features a rendering of a wet-dry vacuum (colloquially known as 'wet vacs', into which users upload WET files by clicking on or dragging files to the black nozzle. Pink text changes between different options every time the tool is visited online and, creatively, summarises the purpose of the tool with a description like 'Sucking the MS BS out of your TEI'.

flows to meet the project's technical requirements – most significantly a method for tracking names and subjects across editions. Developing the project's customised TEI schema followed the same logic as the WET template: utilise the built-in functionality of a tool to create easy-to-use systems for editing XML documents for users who may or may not be familiar with markup languages. Like any XML schema, the PSC customisation (a RELAX NG schema) provides continuous validation feedback whenever someone working on a document is using an XML editing tool like Oxygen. At a baseline level it will, for example, restrict which encoded document structures (aka 'elements') can appear in different parts of an XML document. For more precise and project-specific control, we integrated Schematron rules into the PSC's schema.[12] A rule-based language, Schematron is used with markup to make assertions about the absence, presence or specific arrangement of data signifiers in a document. For example, we created a Schematron rule to display an error message for 'invalid output' when an encoded date was not formatted according to the ISO standard (YYYY-MM-DD). That is, our added Schematron are formatted to emphasise and identify errors in a manner more familiar to users new to editing XML documents.[13]

Over the duration of the PSC's implementation grant (2020–4), this document transformation workflow has undergone many major changes including updating legacy encoding structures from past MHS digital edition projects to metadata elements in the <teiHeader>[14]

12 For more information about Schematron and TEI customisation, see the tutorial on 'XPath and Schematron for TEI Customization' by Syd Bauman for the Women Writers Project seminars (2016): https://wwp.northeastern.edu/outreach/seminars/_current/presentations/schematron/schematron_odd_tutorial.xhtml.

13 For markup languages like XML, validation describes the status of a document conforming to rules and structure of the schema with which it is associated. XML editing software includes functions to automate this process and, as demonstrated in Figure 2, display error messages or descriptions of invalid encoding.

14 The <teiHeader> is the main, root element in the TEI guidelines for document metadata and features a *file description* to describe the bibliographic infor-

and checking for routine consistency in the user interface and XML output from WETVAC. Similarly, the PSC's customised TEI schema has undergone several versions, sometimes leading to inconsistencies in XML files produced with different WETVAC and schema iterations. Those generations of files were subsequently updated with batch XSLT processing and hand encoding. Having the capacity to adapt and refine this technical process *in response* to developing editorial needs as the cooperative progresses has been fundamental towards creating overall project workflows that are stable, consistent, and (in the long run) sustainable. While there are many moving parts of this document transformation workflow, simply put it converts ingested MS Word documents into consistent, valid and well-formed XML documents which editors can further encode and refine as needed.

The team invested time in these innovations because how an edition creates its document transcriptions is of such critical importance. Based on our previous experiences with digital editing, including training documentary editors in the use of various tools, we believe that although WYSIWYG (What You See Is What You Get) editing environments may ease the digital transcription process for some people, these interfaces can keep users too removed from the semantic markup, leading to compromises in the underlying encoding and, therefore, neglecting some of the core benefits of working in the TEI. This is the basis for the cooperative's stipulation that one editor on each edition serve as an XML mediator, bringing together an understanding of the encoding necessary for digital delivery with a firm grasp of their edition's content and goals. This way editorial integrity is best maintained.

---

mation for the electronic document itself, an *encoding description* that relates encoding practices for a project and the electronic document, a *text profile* that contains contextual information like subject headings, and a *revision description* that logs any significant changes or edits to the document. For more information, see https://www.tei-c.org/release/doc/tei-p5-doc/en/html/HD.html.

Conclusion: community and end-user

As editors of the cooperative look to the future, we must consider the needs of our audiences. At the digital edition level, our audience is familiar: the website end user, those readers who will come and interact with our digital editions. This end user has been considered previously by scholars of digital editing,[15] and indeed has been a focus of ours since the beginning of the cooperative. All websites have the standard end user they consider when designing a site, and the cooperative does as well, taking care to discern search terms and details of user interface. In addition to this traditional end user, however, the cooperative must also consider new member editions, who are also users, and who will employ the site to publish their own editions and who are thinking about the methodology of digital editing. Upon joining, each of the member editors made a commitment to providing free access to the editions (no paywalls), and this will be a requirement for all new member editions. From the perspective of the discipline of scholarly editing, our audience consists of those who consider the cooperative as a methodology of digital editing. These two latter end users will be considered below.

In the coming years, we hope to onboard several new editions and grow as capacity allows. Networking with relevant organisations and institutions, we will look for projects that speak to the ages of reform in the long nineteenth century, relate to people of colour or other marginalised populations engaged in reform movements, and complement the current partners so as to expand the discovery and aggregated research possibilities for our user audiences, in terms of both individuals and historical topics. It is important for the PSC

15 Greta Franzini, et al., 'Digital Editions of Text: Surveying User Requirements in the Digital Humanities', *Journal on Computing and Cultural Heritage*, 12.1 (February 2019): 1–23; Krista Stinne Greve Rasmussen, 'Reading or Using a Digital Edition? Reader Roles in Scholarly Editions' in *Digital Scholarly Editing: Theories and Practices*, ed. Matthew James Driscoll and Elena Pierazzo (Cambridge, England: Open Book Publishers, 2016).

to add more projects and voices in the coming years, both to broaden the research capabilities for our website users and to replace the projects that will be cycling off as their editors finish work on those editions. Editors of new projects coming into the cooperative will have the same level of mentoring as the founding editions. The ethos of collaboration will remain. We know from our own personal experience how important it is to have other editors with whom you can connect when a problem or question arises within an edition.

In addition to mentorship, new editors will have the benefit of the documentation policies and editorial guidelines created by the governance board and the publishing systems staff at the MHS. Indeed, this published documentation will be available to all. And this is where we get to our third audience: that of the wider discipline of scholarly editing. The success of a cooperative consisting of scholars who are new to digital publishing will itself be evidence that robust digital publishing for small- to mid-scale editions is achievable and within reach. Publications do not need to be niche or undertaken at great expense with the cooperative system. Our goal is that our PSC model will be a system that is broadly reproducible by other cultural institutions, archival repositories and libraries.

In many ways, our proposed model and resources for cooperative editorial praxis exemplify what Christopher Ohge has identified as 'pragmatic inventions'. Imagining the edition of the future is in and of itself a pragmatic endeavour. As Ohge argues, 'The framework of pragmatism allows editors to embrace and build upon the differences of previous editorial theories, to create new practices and tools, and to embrace technology as a means for publication, discovery, and experimentation.'[16] Indeed, the creation of a new form of edition publishing collaboratively is in line with this pragmatic approach. The cooperative has remixed various tools and approaches to editing with the end goal of providing a pathway

16 Christopher Ohge, *Publishing Scholarly Editions: Archives, Computing and Experience* (Cambridge: Cambridge University Press, 2021), 15.

for publication for small- to mid-size editions. This ensures more voices will be part of the future of editing, and editors will have a structured publication avenue that has been tested and grows the future community of editors. Publishing as a cooperative is a pragmatic and essential approach to the future of small- to mid-scale digital editions.

References

Alexander, B. and Frost Davis, R. 2012. 'Should Liberal Arts Campuses Do Digital Humanities? Process and Products in the Small College World.' In *Debates in the Digital Humanities,* edited by Gold, M.K. University of Minnesota Press.

Boyles, C., Boyles Petersen, A., Landaverde, E. and Dean, R. 2021. 'Postcustodial Praxis: Building Shared Context through Decolonial Archiving.' *Scholarly Editing* 39.

Flanders, J. 2018. 'Building Otherwise.' In *Bodies of Information: Intersectional Feminism and Digital Humanities,* edited by Losh, E. and Wernimont, J. University of Minnesota Press.

Franzini, G. et al. 2019. 'Digital Editions of Text: Surveying User Requirements in the Digital Humanities.' *Journal on Computing and Cultural Heritage* 12 (1): 1–23. https://doi.org/10.1145/3230671.

Klein, J. T. 2015. *Interdisciplining Digital Humanities: Boundary Work in an Emerging Field*. University of Michigan Press.

Meadows, R. D. 2016. 'Building a Sustainable Digital Edition Ecosystem'. Scholarly Communication Institute. https://trianglesci.org/2016/05/18/building-a-sustainable-digital-edition-ecosystem/.

Ohge, C. 2021. *Publishing Scholarly Editions: Archives, Computing and Experience*. Cambridge University Press.

Pannapacker, W. A. 2009. 'The MLA and the Digital Humanities.' *Brainstorm* (blog). *The Chronicle of Higher Education Online*, 28 December.

Rasmussen, K. S. G. 2016. 'Reading or Using a Digital Edition? Reader Roles in Scholarly Editions.' In *Digital Scholarly Editing: Theories and Practices*, edited by Driscoll, M. J. and Pierazzo, E. Open Book Publishers.

Risam, R. 2018. *New Digital Worlds: Postcolonial Digital Humanities in Theory, Praxis and Pedagogy*. Northwestern University Press.

Robinson, P. 2014. 'Digital Humanities: Is Bigger Better?' In *Advancing Digital Humanities*, edited by Arthur, P. L. and Bode, K. Palgrave Macmillan.

Robinson, P. 2017. 'Some Principles for Making Collaborative Scholarly Editions in Digital Form.' *Digital Humanities Quarterly* 11, (2).

Simpson, K. and Ball, H. F. 2021. 'Editing to Avoid Exclusion: Understanding the Subjective Power Dichotomies in Scholarly Editing.' *Scholarly Editing* 39.

Smiley, B. L. 2019. 'From Humanities to Scholarship: Librarians, Labor, and the Digital'. In *Debates in the Digital Humanities 2019*, edited by Gold, M. K. and Klein, L. F. University of Minnesota Press.

Spiro, L. 2012. '"This is Why We Fight": Defining the Values of the Digital Humanities.' In *Debates in Digital Humanities*, edited by Gold, M. K. University of Minnesota Press.

Warwick, C. 2019. '"They Also Serve": What DH Might Learn about Controversy and Service from Disciplinary Analogies.' In *Debates in the Digital Humanities 2019,* edited by Gold, M. K. and Klein, L. F. University of Minnesota Press. https://dhdebates.gc.cuny.edu/projects/debates-in-the-digital-humanities-2019.

Wernimont, J. 2020. *Numbered Lives: Life and Death in Quantum Media*. MIT University Press.