RESEARCH ARTICLE

Founders Online: Early Access: Reflections on Open Access, Crowd Sourcing, and Metadata Standards

William B. Kurtz
University of Virginia, US
wkurtz@virginia.edu

Founders Online, a digital initiative of the National Historical Publications and Records Commission (NHPRC) of the U.S. National Archives, launched in June 2013. Since its debut, the site has attracted over a million visitors interested in learning more about the creation of the United States of America in the words of six of its Founding Fathers. Founders Online contains 177,000 letters or other writings of these men and their contemporaries. Widely used by academics and the general public, the site has demonstrated the value of digital humanities’ emphasis on free access. As a former assistant editor at Documents Compass, a program of the Virginia Foundation of the Humanities, I served as a project manager on the Early Access portion of the project. We worked directly with the staffs of the currently active Founding Fathers documentary editing projects to make preliminary versions of unpublished documents available for early viewing on Founders Online. These Early Access documents will eventually be replaced by fully vetted and annotated versions to be completed later by the documentary editing projects. Relying on a large staff of over thirty people, we transcribed or proofread over 50,000 Early Access documents from 2012 to 2015. My Early Access experience demonstrated the need to give employees constant feedback, to reward them for good work, and to encourage specialization among project staff. My experience also reemphasized the need for unified metadata standards when aggregating different sets of data from multiple projects into a single digital platform.

Keywords: Founders Online; Open Access; transcription; metadata; crowd sourcing; Digital History

Founders Online, une initiative numérique de la National Historical Publications and Records Commission (NHPRC) des Archives nationales des États-Unis, lancée en juin 2013. Depuis ses débuts, le site a attiré plus d’un million de visiteurs intéressés à en apprendre davantage au sujet de la création des États-Unis d’Amérique d’après six des pères fondateurs.
Founders Online contient 177,000 lettres ou autres écrits de ces hommes et de leurs contemporains. Largement utilisé par les universitaires et le public en général, le site a démontré la valeur de l’emphase des humanités numériques sur le libre accès. En tant qu’ancien rédacteur en chef adjoint à Documents Compass, un programme de la Virginia Foundation of the Humanities, j’ai travaillé comme gestionnaire de projet pour la partie d’accès anticipé du projet. Nous avons travaillé directement avec les membres du personnel des projets de montage documentaire de Founding Fathers actifs à l’heure actuelle, pour rendre disponibles en accès anticipé des versions préliminaires de documents non publiés sur Founders Online. Ces documents en accès anticipé seront éventuellement remplacés par des versions entièrement approuvées et annotées qui seront complétées plus tard par les projets de montage documentaire. Comptant sur un personnel nombreux de plus de trente personnes, nous avons transcrit ou relu plus de 50,000 documents d’accès anticipé entre 2012 et 2015. Mon expérience de l’accès anticipé a démontré le besoin de donner aux employés une rétroaction constante, de les récompenser pour leur bon travail, et d’encourager la spécialisation parmi le personnel du projet. Mon expérience a de plus souligné davantage le besoin de normes de métadonnées communes en transposant différents ensembles de données de projets multiples en une plateforme numérique unique.

**Mots-clés:** Founders Online; Libre accès; transcription; métadonnées; externalisation à grande échelle; Histoire numérique

---

Founders Online, a digital initiative of the National Historical Publications and Records Commission (NHPRC) of the United States National Archives, launched in June 2013. Designed and updated by David Sewell and his staff at the University of Virginia Press, the site is an essential resource for scholars of early American history. Since its debut, the site has attracted over a million visitors interested in learning more about the creation of the United States of America in the words of six of its “Founding Fathers”: George Washington, John Adams, Thomas Jefferson, James Madison, Benjamin Franklin, and Alexander Hamilton. With over 177,000 letters or other writings of these men and their contemporaries who corresponded, lived, and worked with them, the site utilizes extensible markup language (XML) encoded documents hosted on a MarkLogic server to provide free, fast, and user-friendly access to the public. Founders Online has received praise from the larger scholarly community, including the Society for History in the Federal Government,
which awarded the site its prestigious Thomas Jefferson Prize for its “outstanding contribution” to scholarship (Ferriero 2015b). This emphasis on reaching a much larger audience than documentary editors have traditionally served, the academic historian, should be considered a model for future digital editions regardless of their scale or subject. As Stephen Pinfield argued, “The major question associated with open access is no longer whether OA [open access] should be at the centre of the mainstream scholarly communication system, but how?” (Pinfield 2015).

As an assistant editor at Documents Compass, a program of the Virginia Foundation of the Humanities that helps documentary editors publish their work digitally, I helped to manage the Early Access portion of the project from 2012 to 2015. Relying on a paid staff of over thirty people, most of whom had little experience in historical or digital work prior to joining our team, we transcribed or proofread over 50,000 Early Access documents. These temporary documents, meant to give the American public an early look at transcripts of historical documents not yet finished by the documentary editing projects, will eventually be replaced by permanent, professionally vetted and annotated versions. Our project was led by Documents Compass Director, Susan H. Perdue, with Assistant Editor Laura K. Baker and myself serving as project managers. The lessons I learned in managing our large staff and in working with a variety of editing projects with different editorial policies and methodological approaches suggest several best practices for other digital humanists considering crowd-sourcing projects utilizing volunteer labor or hoping to impose standardized metadata and workflow procedures when aggregating different projects under one digital roof.

Digital History Meets Documentary Editing
The appeal of publishing humanities work digitally cuts across disciplines and has increasingly become a priority for faculty hiring, cultural institutions hoping to celebrate and grant greater access to their holdings, and for scholars seeking to reach

---

1 For more information on the technical infrastructure that hosts Founders Online, please see Matt Allen, “Founder’s Online: A Lesson in Performance” (Allen 2016) and “Founders Online Launches” (UVA 2013).
new audiences. In the field of history, the importance of writing and producing historical scholarship accessible beyond the walls of academia has become increasingly important even for those who do not consider themselves strictly public historians. In an era of decreasing faculty and research budgets for the humanities, proving that the study of history is valued by the American public is vital for history departments, universities, museums, and other historical entities that rely on public monies to cover their expenses. Certainly projects such as the University of Richmond’s Atlas of the Historical Geography of the United States and American Panorama have demonstrated that digital history can have a wide appeal beyond academia (Stinson 2016).2

The value of publishing online is an important and ongoing debate within the historical subfield of documentary editing. Founding-era documentary editing projects explore the lives of some of the most famous and important figures in United States history. *The Papers of Thomas Jefferson*, begun in 1943 under the direction of renowned editor and historian Julian P. Boyd, is widely considered to be “the first modern historical documentary edition.” This project established standard guidelines and editorial practices that have strongly influenced other documentary editions (Founders Online 2016a). By transcribing and annotating the complete works and correspondence of men such as Jefferson, Adams, and Washington, these editorial projects are engaged in some of the most worthwhile publicly-funded historical work. A variety of private and government funders, most notably the NHPRC and the National Endowment for the Humanities, have provided millions of dollars to make these projects possible and continue to be important funding sources today (Founders Online 2016c).

The tireless efforts of project editors and staff to provide accurate and scholarly editions of these men’s writing, however, are both time consuming and costly. In addition, the high cost of publishing their works in print ensures that only relatively few volumes are produced on a yearly basis. This means that the American public, which has demonstrated time and again its interest in the story of the creation of

2 For a good discussion of the value of digital history, how it differs from traditional scholarship, and how it allows historians to reach wider audiences, see “A Conversation with Digital Historians” (Southern Spaces 2012).
its nation, has limited opportunities to benefit from this important work. While the most important editions did start to go online in the early 21st century, these projects, such as University of Virginia Press’s subscription-based service Founding Era Collection, often existed behind a paywall that limited their use to universities able to pay for access (UVaP Founding Era Collection 2017).

### A Digital Edition Success Story: Founders Online

Founders Online was a congressional initiative to fix this problem (Senate Hearing 2008, 110–334). Proposed in 2008, funded in 2010, and launched three years later in the summer of 2013, Founders Online made freely available online some of the most important editions of the Founding Fathers. The site provides an intuitive search interface and the ability for anyone with an internet connection to access the letters and writings of George Washington, John Adams, Thomas Jefferson, James Madison, Alexander Hamilton, and Benjamin Franklin. The site not only opened up these resources to universities who could not afford Rotunda, but it also made access to the documents that explain the creation of the American Republic easily available to K-12 educators and their students as well as the larger American public. The National Archives underscored its goal of opening these documents to new audiences, particularly to educators, by having students participating in the National History Day contest in Washington, D.C., in 2013 attend the launch of the Founders Online website. These students demonstrated its use to an audience of editors, archivists, politicians, and journalists. Since its launch, the site has attracted over one million visitors both inside and outside of the academic and scholarly history communities (National Archives, 2010, 2011, and 2013a).

Founders Online’s simple site design, with a central search bar resembling Google’s search engine page, is meant to make using the site and its historical documents as user friendly as possible. Designed by the Ivy Group of Charlottesville, Virginia, the website was user tested on several occasions. In addition to Ivy Group’s user testing, students at Bishop O’Connell High School in Arlington, Virginia, also tested the site to assess its user friendliness for K-12 education. The site’s filters allow users to restrict searches by author, recipient, and date, and its recent addition of teaching resources highlighting important documents on specific subjects such
as the Louisiana Purchase, native Americans, and women, are all aimed at helping
the non-specialist explore the site as easily as possible. This emphasis on the non-
academic user sets Founders Online apart from other digital projects (Ivy Group
2012; National Archives 2013b; Founders Online 2016b).

Founders Online has proved useful to scholars and the general public in both
expected and surprising ways. On accepting the Thomas Jefferson Prize, Archivist
of the United States David Ferriero lauded the many ways in which the site had
been used since its launch. He noted that it was cited in articles in the prestigious
historical journal, the American Historical Review, as well as a number of academic
monographs such as Joel Kavorskey's The True Geography of Our Country: Jefferson's
Cartographic Vision (UVaP 2014). Historian Peter S. Onuf's popular massive open
online course (MOOC), "The Age of Jefferson," made extensive use of documents
on Founders Online as part of its required reading material (Age of Jefferson
2017). The Archives' own educational website, docsteach.org, links directly to
Founders Online transcripts, and the National Humanities Center also helped to
create lesson plans utilizing the site's resources for K-12 teachers (DocsTeach 2017;
Ferriero 2015a).

Users beyond historians have utilized letters on Founders Online in other exciting
ways. Anna Berkes, a research librarian at Monticello, has used the website to help
expose fabricated quotations commonly attributed to Thomas Jefferson. The U.S.
Supreme Court cited a document on Founders Online in its 2013 decision in the NLRB
v. Noel Canning case. Perhaps the most creative use of the site was by a composer and
professor at the Berklee College of Music, Gates Thomas, who composed a cantata
based on a letter by George Washington. Perhaps most gratifying has been the
feedback from the many users of the site who have helped to identify a small number
of errors, either in transcription or annotation, that have come to light and are easily
fixed by UVA Press staff thanks to the site's free and digital publication apparatus.
Founders Online thus serves as a model for the digital history and humanities
communities to reach larger audiences and shows the benefits of funding humanities
research (Walters 2015; Supreme Court 2013; With Good Reason 2014).
Early Access: Challenges in Project Management

Founders Online’s emphasis on a user friendly experience is complimented by its creators’ decision to showcase Early Access documents that have not been fully vetted by professional editors. Whereas most projects wait to publish letters only after they have been fully annotated, a decision was made early on that the American public should have access to over 50,000 documents that have yet to be completed by their relevant documentary editing project’s staff. Thus the Early Access project was created to transcribe and proofread these documents, which will be temporary placeholders until professionally edited and annotated versions are completed. These Early Access letters provide insight into important parts of the Founding Fathers lives that the site’s users would have otherwise been without for many years. Transcribing or proofreading these Early Access documents would not have been possible without the NHPRC’s desire to give the American public access to as many documents as possible as quickly as possible (National Archives 2011).

In 2009, Documents Compass received a contract from the U.S. National Archives to complete a pilot project to show whether providing early access to unpublished letters was feasible. Directors Susan Perdue and Holly Shulman, both veterans of the field of documentary editing and leaders in publishing digital editions, proved that the project was indeed possible, leading to a second contract awarded in 2011. The Early Access project officially began on January 2012 and finished in December 2015. The project employed a large staff of proofreaders, sometimes numbering over 30 temporary or graduate student employees, working under the direction of myself and my colleague, Laura Baker. Our proofreaders generally had degrees in the humanities and social sciences but were not experts in 18th century handwriting, history, politics, or culture. By the end of the project, however, they developed an ability to read even the most difficult penmanship and gained a much deeper understanding and appreciation for the Founding Fathers and their times. A number of them gained exceptional technical and editorials skills, which they used to help me reconcile data in our various databases, locate and digitize letters from
microfilm, and conduct basic historical research to try to ensure our transcriptions’ accuracy.¹

Given that our staff was largely inexperienced in editing or historical work, we designed our processes to help ease employees through the various technical aspects of their work. For example, few of them had any experience with editing in XML. We realized that we needed a user-friendly XML software such as Xmetal, a program whose interface closely resembles that of Microsoft Word. We found that the more popular and familiar program, Oxygen, even in its out-of-the-box “Author” view, was too technical and difficult for most of our staff. We also had to train our employees in the unique style guides employed by each project. We created instructions and processes that were simple and understandable in order to produce the best quality transcripts possible while adhering as closely as possible to project specific requirements. Despite our limited time and ambitious goals, varying accessibility of manuscript sources necessary for transcription and proofreading, and the different requirements and methods of each project,⁴ we believe that by the end of our project our employees had improved their speed and accuracy considerably. Our proofreaders benefited from regular in-person and online feedback about their work from our project managers, and, as they gained experience, they became more adept at deciphering late 18th and early 19th century handwriting. They mastered the basics of XML editing and utilized historical techniques such as using other letters to decipher difficult words or using online resources to look up basic information about people, historic places, and past events. Ultimately, our staff produced

---

¹ Here, I would like to recognize Early Access team members Mark Hawking, Jeffrey Diehm, Dena Radley, Jeffrey Zvengrowski, and James Ambuske for their exceptional contributions to the project. Several of our proofreaders were so moved by their experience that they took part in a special episode of With Good Reason, a public radio show featured locally on National Public Radio (NPR), in order to tell the wider community of their newfound appreciation for early American history (With Good Reason 2014).

⁴ For example, our first goal was to proofread and publish 10,000 letters from the Papers of George Washington and Papers of James Madison projects. Both projects have unique style guides of differing complexity and specificity about how to handle superscripts, double punctuation, capitalization, etc. The need to train employees simultaneously in two different style guides greatly increased the difficulty of our work.
preliminary transcripts that will be useful to Founders Online’s users as they wait for the documentary editing projects to complete their work.

Lessons from Early Access

My experience in Founders Online: Early Access not only convinced me of the importance of taking into account a larger, non-academic audience in designing digital humanities projects, but I also believe it has provided several important lessons for digital humanists seeking to use crowd-sourcing, non-specialist volunteer labor as part of their workflow. It is unlikely that a similar project to Early Access with a large, paid staff will happen to be funded again anytime soon. Still, despite its use of paid employees, my experience as an Early Access project manager speaks to the need for constant feedback to project participants and methods of quality control for crowd-sourced projects relying on voluntary and free labor. Projects such as the Library of Virginia’s Virginia Memory do provide a method of limited peer review, but a more robust method of communication between the editorial staff and volunteers, in my experience, is necessary. It is not always helpful to create a one-size list of “do’s and don’ts” for participants, for different volunteers will need more instruction or correction on different points of transcription. We found that constant individual feedback in person, through email, or via our project management software called Basecamp was the best way to ensure the best transcription and proofreading results possible.

I also believe projects need a reward mechanism to acknowledge their best contributors to keep them engaged and motivated to continue working through to a project’s completion. Transcribe Bentham, a very successful crowd-sourced project based out of University College London, does a good job of explaining the benefits for participants and has a “Hall of Fame” where it lists the usernames of everyone who has participated in the project. However, the Hall of Fame does not distinguish

---

5 The Virginia Memory project relies on Transcribe, an Omeka-based transcription plugin, to allow users to review transcriptions done by their peers. Created in part by Roy Rosenzweig Center for History and New Media at George Mason University and the University of Iowa Library, for more information on Transcribe please see “About the Project,” (DIY History 2016).
users by the number of images transcribed, or by the speed or accuracy of their work. Needless to say when your labor force is working for free it is very important that top performers are encouraged and recognized for their exemplary contributions (see “About Us” and “Hall of Fame,” Transcribe Bentham 2016a and 2016b).

We achieved this in Early Access primarily by giving our employees raises, an option admittedly not applicable to crowd-sourced projects. However, we also found non-monetary ways to reward employees. For example, I rewarded my best employees by assigning them particularly interesting writers and batches of correspondence such as Thomas Jefferson’s letters with John and Abigail Adams. I also made a point of acknowledging our best employees’ contributions publicly through Basecamp, or praising them during our training meetings. Many similar opportunities for rewarding excellent work exist for crowd-sourced projects, whether that is using the principles of gamification as seen on popular language-learning programs like Duolingo or the digital crowd sourcing project Old Weather in which the best users are rewarded for their good work with increasingly high nautical ranks for transcribing weather logs from old ship manifests. Highlighting the names of individuals who made significant contributions to the project directly on its website in this way is very important. David Rumsey’s listing of his top ten geo-coders on his Maps Collection site is a perfect example (Rumsey 2017).

Finally, it is a very good idea to encourage specialization. Particular writers’ handwriting or some data sets may be harder to transcribe or record than others. Instead of letting anyone edit or transcribe a document as is done on some Omeka-based sites, volunteer transcribers should have to demonstrate a proficiency in deciphering the difficult handwriting of such historical figures as Abigail Adams or James Monroe before getting full access to these writers’ letters. Not only does this ensure a higher quality of work in the finished product, but we also found that giving access to this important yet difficult subset of letter writers was a useful non-monetary reward that could be copied by others relying on non-expert, volunteer labor.

Founders Online also demonstrates the difficulties inherent in bringing different projects together under one roof that follow different style guides, metadata conventions, and methodology. Differences across projects are inevitable confusing
to users not familiar with the methods of documentary editing. For example, instead of using a single standard convention to indicate that a word or phrase is difficult to read (e.g. surrounding the word with [brackets]), there were several different ways to represent this common editorial issue across different projects. For better or worse, Founders Online did not attempt to standardize these conventions. Similarly, even though the same historical figures appear in different documentary editing projects, the projects have not agreed on a standardized way to spell their names. Hence the Marquis de Lafayette’s long French name appears differently across projects, making a single search for his letters on Founders Online impossible. While adopting a standard methodology or style guide may prove impossible, the adoption of a names authority lists, perhaps by using the Library of Congress’s authority files and adding new names as necessary, would help users. Another Documents Compass project titled People of the Founding Era, a collective biography of the 18th century U.S., has done considerable work to reconcile different naming conventions across projects, and thus might serve as another potential starting point for standardizing metadata across digital projects focused on Early American history.

Conclusion

Founders Online: Early Access was a complex, digital project with a large paid staff that was finished at the end of 2015. Early Access’s successful completion provides a useful model for dealing with the technical issues of large-scale digital humanities work. It shows the need for rewarding hard-working, accurate volunteers while also imposing uniform standards across disparate sets of data. Ultimately, Founders Online’s continued importance will be its emphasis on free access and user friendliness to professional historians as well as the general public. Its emphasis on digital methods of access will prove useful to more traditional scholars studying larger topics of interest to the public, while its emphasis on being useful to the public and thinking about the non-scholarly world will help make digital history even more appealing to potential funders. Such projects should be encouraged in the future, whether they are similar collections of freely available transcriptions related to a particular topic in American history such as the American Civil War, the
Great Depression, or the Civil Rights movement, or large collections of data or other records useful and accessible to researchers, educators, and the general public.

**Competing Interests**
WK was a paid project manager working on Founders Online: Early Access from 2012 to 2015.

**References**


How to cite this article: Kurtz, William B. 2017. “Founders Online: Early Access: Reflections on Open Access, Crowd Sourcing, and Metadata Standards.” Digital Studies/Le champ numérique, 7(1): 2, pp. 1–14, DOI: https://doi.org/10.16995/dscn.277

Submitted: 19 October 2015 Accepted: 31 October 2016 Published: 13 October 2017

Copyright: © 2017 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See http://creativecommons.org/licenses/by/4.0/.

Digital Studies/Le champ numérique is a peer-reviewed open access journal published by Open Library of Humanities.