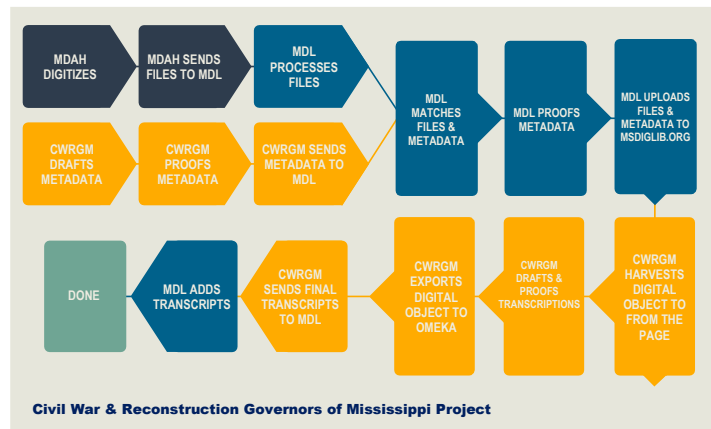


## CWRGM PROJECT METHODOLOGY

The partnership in which Civil War & Reconstruction Governors of Mississippi Project (CWRGM) is grounded allows for focused and steady production represented in this workflow (below). Our partners at the Mississippi Department of Archives & History (MDAH) contribute all digitization work and the Mississippi Digital Library (MDL) provides metadata oversight and hosts digital image files. The Mississippi State University History Department and the University of South Dakota Libraries host project co-directors Susannah J. Ural and Lindsey R. Peterson, who direct the research teams that draft and review metadata, transcriptions, subject tags, and annotations.

**Digitization:** The state's governors' papers are already housed at MDAH, eliminating the need to scour regional and national repositories for these records. As a project partner, MDAH provides all digitization needs, including two-stage quality verification of over 20,000 documents. MDAH transfers TIFF files electronically to the CWRGM partners at MDL in regular batches of roughly 1,000 documents (usually one archival box).



**Metadata:** MDL is an online archival library based at the University of Southern Mississippi that provides digital access to millions of records from repositories across the state. MDL provides CWRGM with digital storage, serving as the repository for electronic surrogates of the digitized documents shared by MDAH. When MDAH sends digital access files to MDL via a secured File Transfer server, MDL assistant director Elizabeth La Beaud and MDL staff convert them to JPEG format and match, name, and folder the files with their metadata.

CWRGM's Senior Assistant Editor, Sarah West, oversees all metadata work (see [CWRGM metadata protocols](#) based on the Dublin Core standard), file matching with metadata when digital files arrive from MDAH, and uploads digitized document files with verified metadata to MDL. West also trains all CWRGM student researchers using our protocols, a customized spreadsheet template and training videos, as well as Zoom training sessions with repeated feedback on work during early training stages followed by weekly meetings. These researchers draft reparative metadata MDAH calendars, updating item descriptions with language that is appropriate in the modern age. In weekly batches, the first reviewer, who cannot also be the person who drafted the metadata (to ensure a fresh set of eyes on the work), analyzes the completed metadata using Google-refine's text facet and Microsoft Excel's pivot tables to identify any inputs that do not conform to MDL's Metadata Best Practices and CWRGM Metadata Protocols. Any concerns are recorded in an issues spreadsheet and returned to the metadata creator for correction. Once all the records in a batch are approved, the metadata

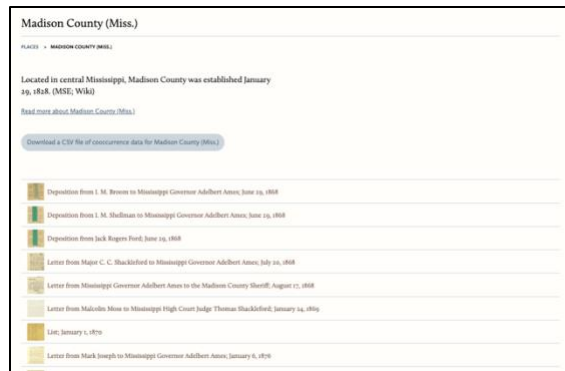
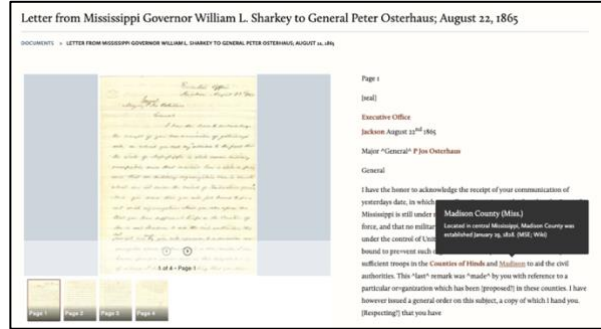
goes on to a second reviewer (different from the first and the creator) who double-checks the metadata for conformity and accuracy, again using Google-refine's text facet and Microsoft Excel's pivot tables. This reviewer also compares the reviewed metadata against the actual digitized document to increase accuracy.

When digitized documents have verified metadata, MDL provides the public with the earliest possible access to these items, earlier than if the project had to wait for transcriptions and enhanced subject tagging features to be drafted and reviewed. MDL and/or the CWRGM GRA uploads new batches of documents as they become available, while the discovery index updates automatically on a nightly basis. But MDL cannot host enrichments such as enhanced subject tagging (discussed below) that dramatically improve document discoverability. That is why CWRGM digital developer Anneliese Dehner designed an Omeka S site for the project that launched with our first 2,000 documents in June 2021. When MDL has 250 more documents at its site than we do at CWRGM.org, we include a link at the Omeka S site called "Sneak Peek" and publicize this via social media to alert users that they can view digital files without transcriptions or subject tagging at MDL. To date, we have not needed to utilize this feature.

**Transcription and Tagging/Enhanced subject tagging features:** Once MDL provides a CSV file for uploading digital document files and verified metadata, CWRGM co-director Peterson uploads these to [CWRGM's home at From the Page](#) (FtP). Volunteers and our team draft transcriptions, and all of these go through a review process *by trained team members only* that follow CWRGM [Transcription and Tagging Protocols](#), with final review conducted by Peterson or Assistant Editor Rachel Brown or Lead Researcher Camp Stewart. As with metadata, all levels of verification are done by different members of the research team. All student researchers go through detailed training with Peterson that includes reviewing protocols and Zoom lessons with regular feedback in their first few weeks followed by weekly meetings with Peterson throughout the semester. This training is grounded in [Documentary Editing literature](#) and in close consultation our protocols.

When CWRGM staff create or review transcriptions, they add enhanced subject tagging features in From the Page that significantly enhance the discoverability of our documents. These tags are tied to an internally controlled vocabulary belonging to the following nine categories: people, places, organizations, businesses, events, occupations, military units, vital statistics, and social identifiers. Most of those categories are self-explanatory; they help users with diverse research interests understand how the collection might be of use to them and how to locate documents of interest. Ural and Peterson created the last category, "social identifiers," to spotlight the experiences of individuals whose voices are often unheard in history, including the enslaved, free and freed people of color, widows, veterans (especially enlisted men), and impoverished individuals. These tags also help users think about new ways to access the collection, suggesting terms like arrest, orphans, impoverished people, and mental illness. Our tagging features remind users of exactly how many documents exist that address diverse topics and offer easy access to them. Available geocoordinates are added in the tagging process for any place, organizations and businesses, and event so these will be ready for future CWRGM mapping projects.

**Annotation:** CWRGM annotates by subject tag (rather than by document) because once an annotation is added to a subject tag, it will populate to every document that includes the term. For example, when CWRGM first annotated the Madison County (Miss.) subject tag, that annotation was populated into the 67 documents then in our collection that contain the subject tag Madison County (Miss) and the annotation appears as a pop-up over the first instance of that term within each document (see right). The annotation also appears on the subject tag's home page with an index of every instance of a document including that subject tag in CWRGM's collection (see below). Anytime future researchers add the subject tag Madison County (Miss.) to a document, that document will automatically include that annotation and be added to the list. This process makes our work highly efficient.



**Selection and Arrangement:** All of the 20,000+ documents of the CWRGM governors' Correspondence and Papers, Military Telegrams, Letters and Petitions, and Executive Journals from 1859-1882 are included in this comprehensive digital documentary edition. In all cases, original documents are digitized to provide high-quality images for scholars and researchers. The only items eliminated from this otherwise

comprehensive project are duplicate documents such as copies of letters and receipts that comprise an estimated one percent of the entire collection. These duplicates are tracked internally in the calendar indexes maintained at MDAH. Because these collections are all government records that are in the public domain, they are free of copyright restrictions/required permissions.

**Publishing Methods:** Today, the [Omeka S home for CWRGM](http://cwr gm.org) (cwr gm.org) has over 10,000 freely available digitized documents with verified metadata, transcriptions, enhanced subject tagging features, and annotations, and it offers users multiple approaches to accessing the collection. Digital developer Anneliese Dehner designed the CWRGM Omeka S site, which allows visitors to conduct [advanced keywords and date range searches](#). Users can also [browse the collection by box and folder arrangement](#), simulating a scholar's research in the archives, including the availability of finding aids. Users can read about and explore the nine subject categories that help non-experts find items of interest, and learn about how to "[Explore the Collection](#)" at the page by the same name. There is also a short essay titled "[How to Get Started](#)" that is linked next to the search bar that appears on every page.

When viewing a document transcription, visitors can click on any tagged word in the transcription to connect to every other document currently at [CWRGM.org](http://CWRGM.org) that contains that term. Tags also provide users with suggested word usage when encountering outdated terms or racial slurs in transcriptions, which are transcribed verbatim from the original document. See, for example, where the cursor hovers over the word "negroes" and the term "Enslaved people" appears. Our suggested terminology varies based on the historical context of the document. That term could refer to "Free people of color," "Freed peoples," or other individuals based on the date and how the term is applied by the author, and it is tagged accordingly. Keyword and most simple indexing do not provide this level of discoverability. As we annotate terms like "Enslaved people" seen in the image here, the text of the annotation will appear and if users click on the phrase, it will take them to a listing of all documents tagged for "Enslaved people." Our site also provides a full index of all documents currently available online. CWRGM maximizes user engagement with a collection that continues to grow in size and in its contributions to scholars and the public.

CWRGM uses Omeka S because it is the leading free, open-source software for online archival and museum exhibitions in digital public history. It has existed for over ten years, is still grant funded, and it has been repeatedly tested and praised by documentary editions upon which CWRGM is modeling itself, including CWGK and JAPP/JADE. CWRGM benefits from working with a developer (Dehner) who has extensive Omeka and Omeka S experience working with digital editions. This site and its enhancements are housed on a Reclaim Hosting server).