Yellow Dust Abode: The Hawley-Suski Letters, 1942-1945

Roxanne Shirazi

Graduate Center, City University of New York
YELLOW DUST ABODE: THE HAWLEY-SUSKI LETTERS, 1942-1945

by

ROXANNE SHIRAZI

A master’s thesis submitted to the Graduate Faculty in Liberal Studies in partial fulfillment of the requirements for the degree of Master of Arts, The City University of New York

2016
YELLOW DUST ABODE: THE HAWLEY-SUSKI LETTERS, 1942-1945

by

ROXANNE SHIRAZI

This manuscript has been read and accepted for the Graduate Faculty in Liberal Studies satisfying the thesis requirement for the degree of Master of Arts.

Stephen Brier

Date

Thesis Advisor

Matthew K. Gold

Date

Executive Officer

THE CITY UNIVERSITY OF NEW YORK
Abstract

YELLOW DUST ABODE: THE HAWLEY-SUSKI LETTERS, 1942-1945

by

Roxanne Shirazi

Advisor: Professor Stephen Brier

On May 6, 1942, Dr. Peter Marie Suski assigned my great-grandfather, Willis M. Hawley, power of attorney to manage his personal affairs when he and his family, along with 120,000 Japanese Americans, were forcibly removed from the Pacific Coast to one of ten U.S. government internment camps. The two men began a regular and frequent correspondence that would continue throughout the war and beyond, producing at least 566 pages of letters between 1942-1960. Dr. Suski and Mr. Hawley wrote as friends, book collectors, and scholars of the Japanese and Chinese languages, and their letters comprise a richly detailed source for biographical and social history. This thesis project uses the TEI schema of XML to encode the letters and identify the topical threads running through them, creating a digital edition that examines the histories of the two men and their correspondence, while considering the ethical implications of digitizing a collection of personal letters.
Preface

On May 6, 1942, Dr. Peter Marie Suski assigned my great-grandfather, Willis M. Hawley, power of attorney to manage his personal affairs when he and his family, along with 120,000 Japanese Americans, were forcibly removed from the Pacific Coast to one of ten U.S. government internment camps. The two men began a regular and frequent correspondence that would continue throughout the war and beyond, producing at least 566 pages of letters between 1942-1960. Dr. Suski and Mr. Hawley wrote as friends, book collectors, and scholars of the Japanese and Chinese languages, and their letters comprise a richly detailed source for biographical and social history. Though densely written and, at times, highly technical in nature, the correspondence contains valuable information about life in the internment camps, the demographic and social transformation of the Little Tokyo neighborhood of Los Angeles, the practice of book collecting, and the spread of texts and artifacts from East Asia to the United States in the years immediately following World War II.

Ever the collector, and with an awareness that the correspondence might prove interesting to future scholars, Mr. Hawley saved the letters he received from Dr. Suski and interfiled them with carbon copies of his own typewritten replies. The collection was kept intact beyond the deaths of the two parties (Dr. Suski in 1960, and Mr. Hawley in 1987), by his granddaughter, Frances Seyssel—my mother—who inherited Mr. Hawley’s estate, including his library and small publishing company, Hawley Publications. By this time, the families, which had once been quite close, had drifted apart and were virtually unknown to each other.

In 2004, a reunion of sorts occurred when Dr. Suski’s great-granddaughter, Lauren Suski, emailed my mother:

My great grandfather is Peter Marie Suski. He wrote My Fifty Years in America. Is the book you are selling in English? Please let me know as I am interested in purchasing it.
My family has several copies of the original book in Japanese. We have wanted to translate it for some time now. Thanks.¹

Lauren Suski had located what she thought was simply the publisher of her great-grandfather’s books, not knowing the long history between the two men. Yet my mother, who had found the letters tucked away on a basement shelf, was elated to discover that the Suski family was still located nearby in the Los Angeles area, and she set to work reviving and sharing those memories which had previously been lost. Lauren Suski and family, including her father, Dennis, and grandfather, Joe, among others, visited the Hawley house that is referenced in the letters, and which Joe himself remembered visiting as a child. Ms. Seyssel shared the correspondence with the Suski family, who had been completely unaware of its existence, and gave Dr. Suski’s originals to them to keep as part of their own family heritage. She made several full copies of both sets at that time.

Ten years later, when I began working with the copied letters for this project, I knew that I would need to obtain copyright permission from Dr. Suski’s descendants in order to digitize them. I also wanted to locate the originals to ascertain whether I could photograph them directly to create an online archive of digital facsimiles. Yet, despite the reunion which had occurred in 2004, the families had once again drifted apart—email addresses changed, some people moved, while others had passed away. Fortunately, the arrival of social media meant that, with some help from Google and Facebook, I found Lauren Suski and reunited the families for the second time.

Lauren recalled meeting my mother and remembered the letters but had never looked closely at them. It turns out that the box of letters had been misplaced among the belongings of Joe Suski, who had died in 2011, and she was not entirely sure where they now were (though she assured me, they were somewhere). It was not until I provided Lauren with a PDF copy of the

correspondence that she was first able to read the words of her great-grandfather, as written to
my great-grandfather.

As she read, I read. I read first with an eye towards uncovering historical details. Then I
read looking for what was left unspoken and unacknowledged. I recognized traces of my own
family history, and found myself becoming uneasy with the private revelations from Dr. Suski,
as though I should not be intruding in these intimacies. But more than anything, I began to
understand the complexity of these two men and their relationship with each other and the world.
These letters are filled with the hopes and realities of two tumultuous lives, lives with which I am
now closely acquainted. Perhaps not surprisingly, the personal nature of this project has both
heightened my investment in it and has caused me to consider more deeply the affective nature
of archival research.

The title of this thesis is drawn from Dr. Suski’s name for his quarters at the Heart
Mountain internment camp, “Yellow Dust Abode.” The text of his explanation for the name is
transcribed below and reproduced in figure 1.

He writes:

I am not sure if I sent you impression of my seal. Guess I did not. Characters in this seal are

heart peak yellow dust place

First I found a white rock soft, very much like chalk. I flattened one side on a grinding
stone. Engraved characters in a modified SW. Didn’t like it, besides the seal looked like a
baby’s foot (see black impression)

So I rubbed on a stone and engraved a grass style of the same characters and turned the
seal top side down. It looks more like a cucumber.

This is the name I have given to the room which faces south near hospital front entrance.
When wind blows in yellow dust, the table top and floor are full of dust. So I named it

‘Heart Mountain yellow dust abode’
Dear Willis;

8/12/45

I received by mail:

1 Wen kuei T'ang catalog
7 pieces of soap stone blank seals
1 piece sandpaper
1 sharpened file
1 samisen plectrum

I am not sure if I sent you impression of my seal. I guess I did not. Characters in this seal are:

心岳黄尘居

First I found a white rock soft, very much like chalk. I flattened one side on a grinding stone. Engraved characters in a modified SW. Didn't like it, besides the seal looked like a baby's foot (see black impression).

So I rubbed on a stone and engraved a grass style of the same characters and turned the seal top side down. It looks more like a cucumber.

This is the name I have given to the room which faces south near hospital front entrance. When wind blows in yellow dust, the table top and floor are full of dust. So I named it:

"Heart Mountain yellow dust abode"

Well so long

P.M. Pete

Figure 1. Source: P.M. Suski to Willis M. Hawley, 12 August 1945. Hawley-Suski Letters. Private collection.
Acknowledgements

This project would not exist were it not for the support and encouragement of many individuals. First, I would like to thank the members of the Suski family, especially Lauren, who generously gave their go-ahead to my exploration of these materials. And to my mother, who found, saved and shared the letters—and the deep family history behind them—with all of us: thank you.

I owe a tremendous debt to the many librarians and digital humanities colleagues I have interacted with on social media and at local workshops, who continue to share their expertise—as well as their data and documentation—through blog posts, tutorials, and, of course, conversations. I would also like to thank the TEI community for providing such detailed instruction and support to individual researchers such as myself, and the team at TAPAS Project for providing a venue through which to publish the encoded letters online.

I would also like to thank my advisor, Dr. Steve Brier, for his incisive comments and unyielding patience as I sputtered along in fits and starts trying to balance academic research and caretaking responsibilities.

Finally, I want to thank my partner and my children for their tireless understanding throughout my time in graduate school, and for allowing me the mental space to explore these ideas while keeping me grounded through life’s frustrations.
Contents

Preface .............................................................................................................................................. v
Acknowledgements ........................................................................................................................ ix
Illustrations .................................................................................................................................... xi
Tables ............................................................................................................................................ xii
Introduction ..................................................................................................................................... 1
Literature Review ............................................................................................................................ 6
Methods........................................................................................................................................... 8
  Presenting TEI Documents Online ............................................................................................. 9
  Creating a TEI P5 Custom Schema ............................................................................................ 11
  Controlled Vocabularies and Terminology ................................................................................ 14
  Named Entity Extraction for Personography and Orgography ................................................. 16
Reflections .................................................................................................................................... 18
  Encoding Absence .................................................................................................................... 18
  Considering full-text resources ................................................................................................. 19
APPENDIX A: TEI Header Template .......................................................................................... 25
Suggested Readings ...................................................................................................................... 29
  TEI Training Materials ............................................................................................................ 29
  Digital Projects ......................................................................................................................... 29
  Literature on the Internment and the Japanese American Experience ...................................... 29
Works Cited .................................................................................................................................. 32
Illustrations

Tables

Table 1. TEI Modules Included in the HSL Schema........................................................................................................ 13
Introduction

The Hawley-Suski Letters Project is a digital scholarly edition of the correspondence between Dr. P. M. Suski and Willis M. Hawley, created using the TEI (Text Encoding Initiative) schema of XML and initiated as a digital humanities master’s thesis. The first phase (that encompassed by this thesis) attempts to create the intellectual and technical infrastructure for an encoding project, focusing on the letters exchanged between 1942-1945, the period in which Dr. Suski was confined at the Santa Anita Assembly Center and Heart Mountain internment camp in Cody, Wyoming. The bulk of the project thus far has been centered on analyzing the content of the letters, creating the customized TEI schema for encoding them, and producing the supplementary files containing biographical information (personography) and organizational information (orgography), which will be linked during the encoding process. The actual number of letters that have been encoded is quite nominal; in this respect, the outcome of this project has been a proof-of-concept to demonstrate my understanding of the technical processes and intellectual considerations that are involved in creating a digital humanities project.

My goal was to create a detailed document structure that would facilitate linking to contextual material in other online sources, as well as to allow connections to be made between the documents within the collection. I wanted to provide readers multiple threads to follow, multiple paths for their own exploration—but guided by my hand as an expert in this collection of letters.

To do this, I first had to immerse myself in the scholarly literature on the Japanese American internment experience and begin an investigation into the histories of the Suski family, as well as my own. These are discussed further in the “Literature Review” section, along with the many sources I consulted in my journey through the TEI guidelines. While studying and
uncovering these histories, I was busy building the technical capacity necessary to undertake a
text encoding project, learning everything from basic XML encoding to file management
practices to establishing the project’s controlled vocabulary and metadata conventions; these
activities are detailed in the “Methods” section. Finally, I’ve included my analysis of the process
of working with these letters as a digital project and consider more deeply issues related to
archives, absence, and historical analysis in the section titled, “Reflections.”

Although I have been trained as a librarian and archivist, I have attempted to approach
this project from the perspective of the historian creating a digital history project. That is, rather
than offer an “archive” of digital sources to users with the expectation that they will make of
them what they wish, I want to offer my interpretation to readers. This is an entirely different
approach than I was taught in library school, where subject expertise is separate from the
technical nature of a project, in part to enable unintended uses by scholars across the disciplines.
While libraries often create online exhibits for their collections, my project resembles more
closely a scholarly edition—as an annotated collection of letters—than the topical explorations
more commonly found in digital content management systems like Omeka. For this reason, I’ve
decided to use the TEI (a markup schema with a long history in literary text encoding) to closely
annotate and link the letters through markup, rather than focus on displaying them visually
online.

I must note, however, the way in which this choice has been received by my colleagues in
the digital humanities community, especially within the narrower field of digital history. There is
nothing flashy or innovative about creating TEI-encoded texts at this point, despite the fact that
there exists a robust scholarly community, an established journal, and international conferences.
Moreover, digital literary studies would appear to have a monopoly on the use of TEI for digital
scholarly work. In 2014, historian Kathryn Tomasek commented on Twitter, “I’m interested in writing about uses of #TEI by #historians. This international standard seems to have dropped below our discipline’s radar.”2 Responses to this tweet, along with earlier similar queries by myself and Michelle Moravec, indicate that while TEI encoding is perhaps useful in teaching environments, the technical barriers to transforming and publishing XML files on the web are sufficiently prohibitive to limit exploration by historians, especially when compared to other techniques that confer more easily accessible—and, perhaps, more immediate—results. This is further compounded when we consider the activities that historians typically seek to perform on digitized documents, such as identifying people, places and organizations or visualizing documents on a map or timeline. These activities are better achieved through alternative techniques of working with machine-readable texts, such as natural language processing and topic modeling. The use of programming techniques to discover connections within or across historical documents is a different approach to digitized texts than boutique encoding projects, which are less exploratory in nature and prioritize scholarly annotation through markup.

Faced with this quandary of how to position my project within the broader scholarly trends of digital history, I began again to explore the notion of a digital scholarly edition and soon discovered the area of work known as documentary editing. Guided by the resources available through the Association of Documentary Editing (ADE), I came to see that the work I wanted to perform did, in fact, have a place in the academy as well as a community of scholars (however niche it may be). Moreover, I came to understand the important distinctions to be made between my own digital edition and the more common “digital critical editions” created by literature scholars. Where textual editors “consciously applied critical judgment and scholarly

experience to produce new, editorially emended texts for their audiences” to “recover words, characters, and phrases of the lost archetype or original,” documentary editors privilege the evidentiary value within a document and perform less editorial intervention in (re)constructing the text in electronic form. As explained in Kline and Perdue’s *A Guide to Documentary Editing* (an ADE publication):

The documentary editor’s goal is not to supply the words or phrases of a vanished archetype but rather to preserve the nuances of a source that has survived the ravages of time. Documentary editing, although noncritical in terms of classical textual scholarship, is hardly an uncritical endeavor. It demands as much intelligence, insight, and hard work as its critical counterpart, combined with a passionate determination to preserve for modern readers the nuances of evidence.³

Understanding this distinction between the critical edition and the documentary edition has helped me situate my thesis project within the practice of scholarly editing; whether or not this activity is counted along the spectrum of digital history, and, more broadly, the digital humanities, is less clear to me now than when I began. Regardless, it is clear that working with the TEI can be valuable as a heuristic activity, especially as an introduction to the way in which textual materials become machine-readable.⁴

For now, I am satisfied that the process of examining the Hawley-Suski letters and deciding what kind of interpretive concepts should be noted, and what subjects and terms should be contained in the taxonomy, is itself a form of scholarship that parallels the act of writing a historical narrative on the subject. That is, I am identifying, evaluating, and extracting the meaningful segments of the source material and arranging them in such a way that the reader—or viewer, or user, or whomever might open up the project—will be brought to certain

---

⁴. For discussion of using TEI as a heuristic in the classroom, see Tomasek et al., “Discipline-Specific Learning and Collaboration in the Wheaton College Digital History Project” and Green, “TEI and Libraries: New Avenues for Digital Literacy?”
conclusions. They will be able to see connections that I’ve identified as well as explore them further right from within the project.

This merging of the technical processes with a disciplinary (or, in this case, interdisciplinary) approach to the materials is the crux of what it means to bring scholarly editions into the digital realm. As Kenneth Price notes, “Mere digitizing produces information; in contrast, scholarly editing produces knowledge.” Further, creating this project as a master’s thesis is an exercise in developing the kind of tacit knowledge that comes with building tools in the context of digital humanities, while using my own reflective capacities and interdisciplinary training to make explicit what has remained implicit in my studies.

Literature Review

In this discussion I use the term “literature” broadly to encompass not only published books and journal articles, but also relevant digital projects, blogs, LISTSERVs, wikis, and other venues of community discussion that commonly address TEI and text encoding.

I was fortunate to discover early on that there is a TEI Special Interest Group (SIG) on Correspondence that is actively working to develop a standard module that would contain elements used to encode correspondence materials. This SIG maintains a wiki (http://wiki.teic.org/index.php/SIG:Correspondence) which lists completed and in-process letter projects, meeting notes and discussion related to their activities, and also has a dedicated LISTSERV (TEI-CORRESP-SIG). Reviewing these materials has greatly influenced the approach to encoding that I have taken with this project.

As background reading, I reviewed several books and articles related to the Japanese American experience with internment, as well as the growing number of digital projects available; these are listed in Suggested Reading. In this exercise, I focused not only on the experience of internment, but also on explorations of community, intergenerational conflict, and the history of racial prejudice and subjugation of Asian Americans.

In addition to the full histories of Heart Mountain provided by Nelson and Mackey, I found the examination of maternity care by McKay and the chapter on the Heart Mountain hospital strike by Fiset to be most valuable in reconstructing the environment in which Dr. Suski worked. Although there is a wealth of literature on the personal experience of internment, I only briefly consulted these accounts due to time constraints; of these, Yoshiko Uchida’s Desert Exile

6. Nelson et al., Heart Mountain; Mackey, Remembering Heart Mountain; Mackey et al., A Matter of Conscience; McKay, The Courage Our Stories Tell; Fiset, “The Heart Mountain Hospital Strike of June 24, 1943.” Of these, only McKay’s work directly mentions Dr. Suski.
and Miné Okubo’s *Citizen 13660* were exemplary for my work locating narratives within Dr. Suski’s letters. Kandice Chuh’s chapter on internment in *Imagining Otherwise* helped me perceive the broader implications of an event in which “racialized identity became so extensively materially significant.”

The Suski family appears in the background of several works on Japanese American community in Los Angeles, with Dr. Suski’s daughter Louise garnering the most scholarly attention as the editor of the first English-language page of a Japanese community newspaper (*Rafu Shimpo*). A recent book, *Nisei Girls*, includes material on Julia and Louise Suski, while Kurashige’s “Biculturalism” discusses Louise’s activities with the Japanese American Citizen’s League (JACL), with some mention of her father’s sensibilities. Two graduate works were treasure troves of information on names and organizations relevant to the Hawley-Suski letters: Hillary Jenks’s Ph.D. dissertation on Japanese American communities in Los Angeles, which includes interviews with Joe Suski (Dr. Suski’s son), and Kaori Hayashi’s master’s thesis on the *Rafu Shimpo*, which includes interview material from Louise Suski.

I also spent considerable time learning about book history and collecting, which enabled me to understand the “shop talk” that Dr. Suski and Mr. Hawley engaged in as book collectors. I was extremely fortunate to have the opportunity to study the history of the book in China with Dr. Soren Edgren at Rare Book School. Dr. Edgren was a (much younger) colleague of Mr. Hawley and he shared his personal recollections as well as his subject expertise on the contents of the Hawley library.

---

8. Kurashige, “The Problem of Biculturalism.” It was Kurashige’s article that directed me to the archives of Louise’s work with the Japanese American Evacuation and Resettlement Study, including her 1945 biographical sketch of Dr. Suski (see Suski, “Biography of Father”).
Methods

The research process for this thesis project can be divided into four stages:

1. Building technical capacity: learning principles of XML and TEI, including how to develop a text encoding project from start to finish.

2. Literature review: read background materials on Japanese American internment and conduct an environmental scan of complementary digital projects.

3. Customizing metadata and standards through subject analysis: determine what TEI P5 modules and elements will be included in my custom schema based on the information gleaned through stages 1 and 2; create a controlled vocabulary (taxonomy) for subject terms based on a modified LCSH and Denshō Digital Repository Controlled Vocabulary.

4. Encoding the texts: create the TEI documents using the Oxygen XML editor, validating against the custom schema I have created in stage 3. Create accompanying personography and orgography files in XML (TEI) format.

5. Presenting materials online: upload sample project documents to TAPAS Project to test implementation of my custom schema.

It is important to note that my project differs from ordinary cataloging and classification activities—instead of applying multiple subject terms to the entire document, I assigned concept keywords (using my controlled vocabulary, drawn from a modified LCSH and Denshō Digital Repository Controlled Vocabulary) to segments of the letters. These classified segments form the “topical threads” that I have chosen to make visible in the final presentation of the project. Encoding the letters at this level of detail will allow me (in a future project stage) to develop a system for querying the letters based on the threads identified in them. Importantly, though, the
fact that the letters have been encoded in XML (at any level) means that they can be put to use in a variety of ways in the future, not just for my own digital project. For example, they could be presented as a traditional digital archive, or analyzed using any number of visualization tools.

Presenting TEI Documents Online

One of the biggest challenges of this project has been determining how to present my encoded letters online. Much of the library literature assumes access to digital asset management systems such as Fedora, which are not appropriate for my “boutique” project. Another typical workflow would be to manage the TEI-encoded XML files using an XML database such as eXist and running XSLT transformations to render the files and queries as web pages. This, however, requires a level of technical knowledge that is out of the scope of my project as a one-semester master’s thesis. Recently, the TEICHI Framework has been developed to manage TEI Lite documents within the Drupal content management system. However, the TEI Lite specification does not address specific features of correspondence, and I lack technical knowledge for developing and utilizing Drupal, which I understand carries a steeper learning curve than I could accommodate within this thesis project.

My initial thesis proposal included a provision for an Omeka site and using the TEI Boilerplate (http://teiboilerplate.org/), developed by John Walsh at Indiana University, to present the encoded letters. The TEI Boilerplate is a simple solution to the problem of publishing the TEI documents, as it instructs web browsers how to render the XML so that one does not have to delve into the complicated XSLT language that would ordinarily be necessary to transform and
display the encoded documents.\footnote{Walsh, “TEI Boilerplate.”} It works with Omeka, provided that the files are housed on the same server as the Omeka installation.

However, I decided to forego the idea of an Omeka component for this project, for multiple reasons. First, my limited resources and extenuating circumstances regarding the source material made it difficult to obtain high-resolution scans of the original letters, which collectively span 566 pages across the entire collection. While this remains a long-term goal for the project, it simply was not feasible to undertake such an effort as part of the thesis project. Second, after spending sufficient time with the TEI guidelines I realized that the bulk of my time would have to be spent developing my custom schema and TEI header for the project, which meant that the actual number of encoded letters I could produce on a thesis timeline would be rather limited. Had I decided to go a different route with this project—for example, by processing the digitized letters through hand-corrected OCR (optical character recognition) and analyzing them with a tool like Palladio (\url{http://palladio.designhumanities.org/#/}) —it would have made sense to invest in high-quality digital images. But because my primary goals were to approach this as a scholarly editing project, and hand-correcting OCR of mediocre scans would be almost as time consuming as manually transcribing the documents, I opted to focus on encoding the letters first and leave the digital facsimiles for a later stage.

Fortunately, there was yet another simple solution for displaying my TEI-encoded files online without crossing into the complexities of XSLT. TAPAS (TEI Archiving, Publishing, and Access Service, \url{http://beta.tapasproject.org/}) is a service that is designed for projects similar to mine—student work, and other small encoding projects that lack institutional resources for
publishing TEI documents. Users create an account and upload their XML files, and the site provides three views of the uploaded documents: rendered with the TAPAS stylesheet, rendered with the TEI Boilerplate stylesheet, or the raw XML. In February 2014, it began beta-testing and, though it is still under development, I have selected it for the Hawley-Suski letters because it provides long-term data storage and curation and has an active user community with expertise in the TEI. In considering the continuing life of the project after the thesis portion is deposited, I feel that working with the TAPAS service is the preferred solution.

Creating a TEI P5 Custom Schema

The use of TEI to encode correspondence has been limited, but there is a growing movement to establish standard conventions for working with letters and creating standards for interoperability across collections. In the latest version of TEI (P5), a new set of elements was added to capture information about correspondence as an event, outside of the material and structural considerations that were already available within the manuscript module. These elements, contained within the `<correspDesc>` (correspondence description) element, appear in the `<profileDesc>` (profile description) section of the TEI header area and are meant to convey the context of a piece of correspondence by indicating the participants and related communications. The first, `<correspAction>`, includes @type values such as “sent,” “received,” “transmitted,” “redirected,” and “forwarded,” which indicates the extent to which the TEI can accommodate complicated communication structures and a variety of contextual participants.

10. “About | TAPAS Project.”
11. Stadler and Seifert, “TEI-Correspondence-SIG/correspDesc.”
The second important element for encoding correspondence is `<correspContext>`, which uses the TEI’s reference and pointer structures to refer to “previous” or “next” correspondence. The discussions occurring on the TEI-CORRESP-SIG list and on the TEI Wiki have been helpful in determining what modules to include in my customized schema and, more generally, how to approach the encoding of the letters.

That said, the process of creating the customization is both more important and more time-consuming than I had anticipated. The TEI P5 standard contains 505 individual elements that are grouped into 21 modules, and it is necessary to define a TEI schema prior to encoding any documents. One can either select a pre-determined schema, or customize one by selecting relevant modules and/or individual elements and including them in a custom schema. The online tool known as Roma facilitates this process, and it generates both the “blueprint” for the schema (known as an ODD file in TEI) as well as the schema itself.

To create a custom schema for the Hawley-Suski letters, I first read through the entire TEI Guidelines to better understand the options available, and the different levels of encoding that could be implemented. I then reviewed the work of the TEI-CORRESP-SIG to consider what letter-specific information I wanted to draw out of these documents. The modules included in hsl_tei.xml are listed in Table 1 along with the primary justification for its inclusion.
Table 1. TEI modules included in the HSL schema

<table>
<thead>
<tr>
<th>Module Name</th>
<th>Formal Public Identifier</th>
<th>Justification for Inclusion in hsl_tei</th>
</tr>
</thead>
<tbody>
<tr>
<td>header</td>
<td>Common Metadata</td>
<td>Required; contains all contextual metadata for the collection and participants.</td>
</tr>
<tr>
<td>linking</td>
<td>Linking, Segmentation, and Alignment</td>
<td>Required; facilitates references between letters.</td>
</tr>
<tr>
<td>core</td>
<td>Common Core</td>
<td>Required; contains elements common to all TEI documents.</td>
</tr>
<tr>
<td>TEI</td>
<td>TEI Infrastructure</td>
<td>Required; defines classes used in all other modules.</td>
</tr>
<tr>
<td>textstructure</td>
<td>Default Text Structure</td>
<td>Required; contains the elements forming a TEI document, including the TEI header element.</td>
</tr>
<tr>
<td>namesdates</td>
<td>Names, Dates, People and Places</td>
<td>Producing a detailed personography is a key part of this project and this module provides the ability to refer to detailed personal, organizational, and geographical names.</td>
</tr>
<tr>
<td>transcr</td>
<td>Transcription of Primary Sources</td>
<td>Allows for accurate description of the features of archival documents.</td>
</tr>
<tr>
<td>figures</td>
<td>Tables, Formulae, Figures</td>
<td>There are several drawings, lists, and a few tables present in the letters and future stages of the project may find these elements to be of value.</td>
</tr>
<tr>
<td>msdescription</td>
<td>Manuscript Description</td>
<td>Enables material aspects of the documents to be drawn out in the encoding process.</td>
</tr>
<tr>
<td>analysis</td>
<td>Analysis and Interpretation</td>
<td>Provides the ability to associate particular “spans” of text, including below the word level, with interpretations.</td>
</tr>
<tr>
<td>tagdocs</td>
<td>Documentation Elements</td>
<td>This will be necessary to generate documentation of the custom schema in future stages of the project.</td>
</tr>
</tbody>
</table>

The next step was to construct the TEI header. This component of a TEI document is more formally referred to as “Common Metadata” module, as it contains the prefatory material about the encoded text, including details of the encoding process, the source documents, and
contextual information. Although it is possible to enumerate personography information and controlled vocabularies directly within the TEI header, I chose to create external files and simply link to them from the header. Part of this decision stems from my earlier decision regarding the document structure of the project: for the purposes of this thesis I considered creating a single TEI document containing the header and a handful of letters in the <text> element, divided by milestone breaks. However, since I plan on continuing the encoding project beyond the life of this thesis, I chose to create individual TEI documents for each letter. This requires the somewhat more complicated, yet more flexible in the long run, practice of maintaining the personography, vocabularies, and other referential data in external files and linking to them using the @ref attribute.

To facilitate the production of a separate TEI document for each letter, I created a TEI header template for the project, including areas that were commented with instructions for the encoder. The full TEI header template is reproduced in Appendix A.

Controlled Vocabularies and Terminology

Perhaps the most important intellectual exercise in my project has been to determine the controlled vocabulary to use for my interpretive classifications (the subject “indexing” that allows me to link topics throughout the collection of letters). The Hawley-Suski letters uses a combination of two existing controlled vocabularies: the Denshō Digital Repository’s Controlled Vocabulary (DDR), and the Library of Congress Subject Headings (LCSH). The former contain specialized terms that are more appropriate for the content of the letters, while the latter are

13. I benefitted greatly in this activity from the widespread availability of the XML files for encoded texts online. My TEI header template was modeled in part on the Indiana University Libraries’ TEI header template for the Women Writers Project.
available as URIs and will allow for future linking to outside sources. While the DDR vocabulary terms were immensely helpful in dealing with the subjects relating to Japanese American history, a large portion of what needs to be pulled out of the letters are more generic topics that necessitate also relying on LCSH.

One important issue to be aware of when studying the Japanese American internment experience is the appropriate use of language to avoid historical euphemisms for the forced incarceration of U.S. citizens and legal residents based on their racial identity. The Civil Liberties Public Education Fund (CLPEF), which in the 1990s administered hundreds of federal grants to scholarly projects on Japanese American internment, issued a Resolution Regarding Terminology that encouraged grant applicants to avoid terms such as “evacuation,” “relocation,” and “assembly center,” preferring instead “exclusion,” “incarceration,” and “detention center.”

This is an important consideration, as even in this project I was unconsciously influenced by the language used in the primary sources I consulted; it took a deliberate effort to avoid repeating these euphemisms, and became a large part of how I formed the controlled vocabulary for the digital project. The unsatisfactory state of the LCSH terms stems from the fact that they are quite literally drawn from the literature and changes happen slowly. According to Geoff Froh at Denshō.org, their specialized vocabulary was created when they deemed LCSH to be inappropriate to the goals of their project, in part due to its use of this problematic language: “We felt that LCSH contains some anachronistic terminology, adopted directly from wartime government euphemisms designed to justify incarceration, that is no longer accepted as valid by mainstream scholarship.” Indeed, the standard LC subject heading for materials on the Japanese American internment is “Japanese Americans – Evacuation and Relocation, 1942-

15. Geoff Froh, e-mail message to author, April 22, 2015.
1945.” Notably, however, the standard LC subject heading for the camps themselves are equally anachronistic, but to a different effect: “World War, 1939-1945 – Concentration Camps – Wyoming.” This generic heading is sub-divided geographically, so one could achieve a running list of United States “concentration camps,” arranged alphabetically by state—a list which would surely outrage mainstream politicians today.

Named Entity Extraction for Personography and Orgography

To expedite the process of identifying names and, later, places, that should be included in the personography and orgography files, I performed some lightweight named entity recognition on the uncorrected OCR output of the scanned documents. Although I planned to transcribe the letters manually and identify the names and locations to tag during the encoding process, I wanted to be able to quickly capture a handful of “starter” names to work with at this early stage. Otherwise, the learning exercise of creating a personography and orgography file would be practically meaningless, as the files would be virtually empty until I could encode a large number of letters—which would not ultimately occur as part of the thesis.

To do this, I ran the master PDF file of low-quality scans that I had initially obtained from the Hawley estate through the ABBYY FineReader program to produce an OCR’ed version of the documents. I then saved the output as a plain text file and loaded it into the Stanford Named Entity Recognizer (NER), using the 3-class model to identify locations, persons, and organizations.16 I then extracted the results into a CSV file and cleaned the data using OpenRefine. Once I had a reasonably clean spreadsheet to work with, I exported the results to

16. Michelle Moravec has shared a wonderfully useful tutorial for performing this work on her blog. See “How to Use Stanford’s NER and Extract Results,” History in the City (blog), June 28, 2014, http://historyinthecity.blogspot.ca/2014/06/how-to-use-stanfords-ner-and-extract.html
Excel, so that I could manually add information as I encountered it during my background research and while encoding. I wanted the data in a spreadsheet format so it would be human-readable during the analysis stage, but easily transformed into a basic XML structure when it came time to create the TEI-encoded personography and orgography files.

To manage the TEI encoding, I purchased an academic license for the Oxygen XML editor, and set to work customizing my workspace and learning the software. Though I have only included a handful of encoded letters in this thesis project, the framework for the project that I created—the TEI header template, the personography and orgography files, and the controlled vocabulary—represents a significant part of the effort involved in any digital scholarly edition.
Reflections

Encoding Absence

Much of text encoding work can be construed as akin to Geertzian thick description, adding layers of context and interpretation to a text to uncover hidden connections within and without the documents. The encoder’s task is to annotate, to literally mark up the document with additional information about what is presented in the source material. With the Hawley-Suski letters, however, I struggled with what was missing. Major events that occurred at the Heart Mountain internment camp, such as the hospital strike and the draft resistance movement (the Fair Play Committee) are never explicitly mentioned in Dr. Suski’s letters. Instead, these periods are simply marked by less frequent communication. I wanted to be able to mark places in the letters where events and ideas are not mentioned. But this assumes there is a moment, a phrase, an inkblot, a smudge on the paper to indicate some presence of an absence. In some cases there were such indicators. Willis’s laments about the infrequency of Dr. Suski’s letters in November 1942 (leading up to the hospital strike) and in February 1944 (the peak of the draft resistance movement) could be highlighted and encoded with interpretive tags noting the way in which these events took an unspoken toll. Yet does this really count as absence, or are these the usual archival traces that call out to the historian?

How do I add a layer of what’s truly missing? How can I tag Dr. Suski’s wife and children, who so rarely appear in the letters? What happens when you look at a document for what’s not there, and what span of text should be wrapped in this “absent description”? Perhaps one could generate a list of anti-keywords and anti-tags to indicate what isn’t present. Theoretically, one could tag arbitrary points in the texts, as one already does when creating placeholder milestones to get around the limitations of TEI’s hierarchical structure. Yet I suspect
that the processing instructions for such annotations would become overly complicated and unlikely to transfer across systems. In general, the process of encoding these few sample letters has demonstrated the enormous flexibility of the TEI schema for the purposes of scholarly editing, but I continue to have questions around the potential for text encoding to approach what is perhaps best addressed discursively.

Considering full-text resources

The correspondence between Dr. Suski and Mr. Hawley evokes several themes that are present in the historical literature around Japanese American internment, but I’d like to examine the effect that my personal ties to these letters have had on my approach to the material.

The circumstances surrounding these letters affect not only the content, but the way their sentiments are conveyed to each party. Dr. Suski and Mr. Hawley were very close friends, and despite the appearance of these letters as a professional correspondence around their research and publication activities, there are, throughout, glimpses of the way that these two men began to lean on each other, to depend on each other not just as colleagues but as confidants. Despite an initial awareness of the likelihood of censorship, over the course of their correspondence Dr. Suski and Mr. Hawley show an ease and frankness with each other that indicates the closeness of these two men. And though it was Dr. Suski who himself suggests in February, 1943, that one day the letters might be of interest to others — in some distant future — the conditions in which the letters were written necessitated a sort of honesty that one might not find otherwise. First, there was no other means through which to communicate once Dr. Suski was moved from the Santa Anita Assembly Center, which had been close enough to Hollywood for occasional visits from Mr. Hawley. And due to the business transactions that Mr. Hawley was charged with, not to
mention the scholarly projects that were initiated during this time, they needed to communicate — to be frank with each other and get things done: there were bank deposit boxes to be closed, mortgage payments to be made, checks to be sent and messages to be conveyed. Second, there is ample evidence within the letters that Dr. Suski and Mr. Hawley were very close. Here are two friends who share a common interest that very few others care about, and who have no choice but to communicate solely through the written word. Naturally, the letters are dense with minutiae about any number of topics, personal and professional. These two men wrote with such frequency, and such honesty with each other, that the letters themselves start to cross over into a sort of diary, wherein each author contributes updates on his life events, his family, his work, his thoughts on current events, his eating habits, his material needs, and his future plans.

If one can manage to dig through the administrative detritus of this correspondence, one finds not just a touching tale of friendship, but a historical source rich with details of textual transmission, amateur printing practices, and the numerous other specialized subjects that traverse this collection.

Yet there are also risks associated with this work. After spending hundreds of hours with these letters and the background materials related to them, after inviting these two lives into my mental space and getting to know them so intimately — how am I to do the work of the historian and share these personal details that I’ve discovered — details that have now become a part of my life, of my family history? I am torn between wanting to share these rich documents and their lessons with the world, and wanting to guard them, to protect them from false interpretations, from being put to uses that I cannot control.

17. After Dr. Suski left the Santa Anita Assembly Center, for the period under discussion in this thesis, he and Mr. Hawley make no reference to any communications with each other outside of their letters and postcards. After the war, when Dr. Suski was settled in Denver, they were able to talk a few times via short-wave radio.
In Otto Preminger’s 1944 film noir, *Laura*, Detective McPherson (Dana Andrews) famously allows himself to fall in love with the woman whose murder he is investigating, by virtue of his immersion in the traces of her life. His conjecture of what she was like, which has been pieced together by examining her possessions and intimate circumstances through interviews with her friends, is not unlike the historian’s work in the archives. But at what point do I turn from discovering to protecting? That false binary of hero investigator and crooked cop, the thin line of corruption that presents itself when faced with the power of storytelling, is ever-present with the historian’s craft. With the power to reveal comes also the power to conceal.

This, then, forms the guiding question for my digital work with the Hawley-Suski letters: what needs to be drawn out, and what, if anything, should remain hidden? Should they, ultimately, be made available as full text documents online? For despite all of the potential that this collection holds in terms of scholarship, there are issues of privacy and personal relationships that are at stake in opening up these letters. Fortunately, there is already discussion of facilitating a *selective* privacy in projects like Mukurtu, which allows granular permissions and access metadata to be applied to materials, in keeping with indigenous customs and social rules. Perhaps we need to apply these sorts of concepts to more modern materials that have thus far been protected from privacy concerns if only because of the restrictions of copyright. What responsibility to our archival subjects do we as historians (and archivists) hold?

In many ways this project has been about coming to terms with my discomfort with their words. There have been numerous occasions where I’ve questioned my decision to work with these documents, stemming from the ethical responsibility I feel towards the memory of my great-grandfather and my current personal ties with Dr. Suski’s descendants. Yet there are also political and moral considerations that have made me question the utility of bringing these
documents to light. For one, Dr. Suski represents an outlier as far as Japanese American internment experiences are concerned. As a first-generation immigrant from Japan, or Issei, he was not alone in his apolitical stance while in the camps. Yet, because he was a medical doctor and worked in the camp hospital at Heart Mountain, he was usually overworked — a stark contrast to the “little to do and much time to brood” experience noted by Bill Hosokawa.18 As Suski later writes in his autobiographical pamphlet:

I lived within the hospital and gave 24 hour service whenever necessary for the entire period the hospital was operating. Teaming with Dr. Nakaya, I did a number of major surgical operations. I also attended 288 cases of childbirth out of a total of 527 attended by seven doctors. My services at the hospital gave me a very agreeable chance of working with all my might without ever thinking of money matters. Those were really the happiest days of my life.19

What are the personal and political implications of releasing such a statement? How might those words used in reference to the Japanese American Internment experience—“the happiest days of my life”—be misappropriated by others if released without the historical and personal context surrounding it?

One of the core tenets of archival practice is preserving context; for archivists, this frequently comes down to issues around provenance and authenticity. Interpretation is left to the researcher, while archivists strive to plan for unimagined and unintended uses. Digital history projects, on the other hand, tend to serve the interests of the scholars creating them. It is precisely this problem that Lauren Klein describes in “The Image of Absence,” when she attempts to uncover the traces of James Hemings in the Papers of Thomas Jefferson Digital Edition.20 The editors of that project (trained as humanities scholars, not librarians or archivists) had included access points for letter author but all other named references could only be found through a

19. Suski, My Fifty Years in America, 42.
keyword search. This is how they had anticipated users navigating the collection. Most of the information that Klein was seeking, however, was available in the data she received precisely because the letters had been encoded in TEI. That is, TEI “bakes in” the context and interpretation of textual materials right into the data file. So while Klein had to perform various feats of programming and extracting alongside close reading of the texts themselves, it was as though the scholarly work that had already been done by the project editors had been digitally engraved into the files she received.

This ability of the TEI to facilitate thick contextual layers of textual materials is the primary reason why I chose the technique for the Hawley-Suski letters. My concerns around releasing full-text copies of the letters (and all of the sensitive issues contained within them) are somewhat mitigated knowing that the materials are quite literally wrapped in the biographical and historical context from which they came. So while TEI is one of the more traditional techniques for digital humanities research and is hardly considered innovative in digital library circles, I have come to appreciate its unique ability to achieve my scholarly aims for this project.

Tara McPherson, in discussing the multimodal writing platform Scalar, explains:

Working in Scalar, scholars are pulling sets of visual materials from digitized collections into a Scalar project or “book,” encouraging the project’s reader to examine these materials in their own right while also engaging a scholar’s analysis of the materials. Such a project is neither solely a book nor solely an archive, but rather a hybrid space between the two that blends scholarly analysis with a rich trove of primary materials.21

It is clear to me now that the “hybrid space” to which McPherson refers is also instantiated in TEI-encoded texts, but at the level of the data file, not the interface. And bringing scholars to the level of the data file generates a level of digital literacy—an understanding of file structures, metadata, and, more generally, how computers process texts. Working with the

Hawley-Suski letters in this thesis project has shown me that the ability to appreciate highly specialized TEI projects as “neither solely a book nor solely an archive” depends only on one’s ability to read the data itself.
APPENDIX A:
TEI Header Template

```xml
<?xml version="1.0" encoding="UTF-8"?>
<?xml-model href="file:/Users/roxanneshirazi/Dropbox/Hawley-Suski/tei_hsl.rnc"
type="application/relax-ng-compact-syntax"?>
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <!--Remove $ signs after you add the values or content; $ signifies
  that the encoder needs to supply info.-->
  <teiHeader>
    <fileDesc>
      <titleStmt>
        <!--Title of the manuscript; follow pattern "Letter from X to Y, date"-->
        <title>Letter from <persName ref=""></persName> to <persName>$receiver</persName>, <date when="yyyy-mm-dd">$date.</date></title>
        <author ref=""></author>
        <editor ref="">Roxanne Shirazi<editor>
        <respStmt>
          <resp>Transcription and encoding by</resp>
          <persName ref="">Roxanne Shirazi</persName>
        </respStmt>
      </titleStmt>
      <editionStmt>
      </editionStmt>
      <publicationStmt>
        <authority>Roxanne Shirazi</authority>
        <pubPlace>New York, NY</pubPlace>
        <date>2015</date>
        <availability status="restricted">
          <p>Copyright 2016, Roxanne Shirazi.</p>
          <p>Letters reproduced by courtesy of the Willis M. Hawley and P.M. Suski estates. Information contained in this document is provided for non-commercial, personal, or research use only. All other use, including but not limited to commercial or scholarly reproductions, redistribution, or publication, without prior written permission of the copyright holder is strictly prohibited.</p>
        </availability>
      </publicationStmt>
      <seriesStmt>
        <title level="u" type="thesis">Yellow Dust Abode: The Hawley-Suski Letters, 1942-1945</title>
        <editor>Roxanne Shirazi</editor>
      </seriesStmt>
      <sourceDesc>
        <msDesc>
          <msIdentifier>25</msIdentifier>
        </msDesc>
      </sourceDesc>
    </fileDesc>
  </teiHeader>
</TEI>
```
In this letter, this is a typescript letter. Saved and collated with preceding and subsequent letters by Willis M. Hawley. Passed on to his granddaughter, Frances Seyssel. This letter was encoded as part of "Yellow Dust Abode: The Hawley-Suski Letters, 1942-1945," a master's thesis by Roxanne Shirazi, in partial fulfillment of the requirements of the Liberal Studies program at the Graduate Center, CUNY. Only romanized forms of the Chinese or Japanese languages have been transcribed. Where character forms such as kanji are present in the original manuscript, they have been omitted from the transcription and their place marked by a GAP element. The Hawley-Suski letters were encoded with the primary aim of facilitating linking within the letters and between outside sources. To this end, effort was made to manually identify named entities within the text and provide contextual information. Where non-standard spelling is used, both the author's spelling and the standard Oxford English Dictionary spelling has been encoded to enhance search results.
<classDecl><taxonomy xml:id="hslDensho">
  <bibl>Densho Topics Vocabulary, modified for the Hawley-Suski Letters project.</bibl>
</taxonomy></classDecl>

<schemaSpec ident="tei_hsl" start="TEI teiCorpus" prefix="" docLang="en">
  <!-- required modules -->
  <moduleRef key="header"/>
  <moduleRef key="linking"/>
  <moduleRef key="core"/>
  <moduleRef key="tei"/>
  <moduleRef key="textstructure"/>
  <!-- optional modules -->
  <moduleRef key="namesdates"/>
  <moduleRef key="transcr"/>
  <moduleRef key="figures"/>
  <moduleRef key="msdescription"/>
  <moduleRef key="analysis" except=""/>
  <moduleRef key="tagdocs" except=""/>
</schemaSpec>
</encodingDesc>

<profileDesc>
  <!-- Note here the approximate percentage of English vs Chinese and Japanese -->
  <langUsage>
    <language ident="en-US" usage="">American English</language>
    <language ident="zh" usage="">Chinese</language>
    <language ident="jp" usage="">Japanese</language>
  </langUsage>
  <!-- Identifies whether typewriting or handwriting is present, or both -->
  <handNotes>
    <handNote xml:id="H1" script="print" medium="typewriter">Indicates portions of the document composed with a typewriter.</handNote>
    <handNote xml:id="H2" script="cursive" medium="pencil">Indicates autograph portions of the document.</handNote>
  </handNotes>
  <!-- Insert keywords classifying the letter as a whole. -->
  <textClass>
    <keywords scheme="#hslDensho">
      <term></term>
    </keywords>
  </textClass>
  <correspDesc>
    <correspAction type="sent">
      <!-- Insert sender information for the *event* not transcription -->
     <name sameAs="#author"/>
    </correspAction>
  </correspDesc>
</profileDesc>
Suggested Readings

TEI Training Materials

Dalmau, Michelle, and John A. Walsh. “TEI Workshop Information.” *School of Library and Information Science, Indiana University*. http://dcl.slis.indiana.edu/teiworkshop/


Digital Projects


Literature on the Internment and the Japanese American Experience


Works Cited


