Data Management Plan: Implementation Report
Dr. Kara Cooney | Coffin Reuse Research

Prepared by
Savannah Lake, Ashton Prigge, and Marisa Purcell
IS 262B | Dr. Jillian Wallis
June 12, 2019

Table of Contents

I. Overview

Dr. Kara Cooney is a professor of Egyptian art and architecture at UCLA, as well as the chair of the Department of Near Eastern Languages & Cultures. Specializing in craft production, coffin studies, and economies in Ancient Egypt, Dr. Cooney has extensively researched Ancient Egyptian funerary practices, contextual architecture, funerary arts, and material culture.

Her current research explores coffin reuse in Ancient Egypt. For Ancient Egyptians, coffins were an integral part of the afterlife, facilitating the transformation of the dead. However, during the 21st Dynasty (circa 1150 BCE), Egypt and its neighboring civilizations in the Mediterranean and the Near East experienced socioeconomic and political instability, afflicted with drought, famine, and foreign invasion.[1] As a result, trade networks collapsed, and Egypt no longer had access to the materials necessary to create coffins. As coffins were not believed to protect the dead in the long term (instead seen as enabling rebirth immediately after death), Ancient Egyptians resorted to reusing older coffins in order to ensure safe passage to the afterlife for the recently deceased.[2] Dr. Cooney's research revolves around evaluating these coffins for signs of reuse, looking for modifications in decoration, names, and coffin parts.

Dr. Cooney's research on coffin reuse is an expansive endeavor, involving eight years of data collection. All in all, Dr. Cooney has examined over 300 coffins in over 20 museums and private collections around the world, generating over 100,000 photographs as well as qualitative data on each coffin. This qualitative data was collected and saved as PDF field note files, which were then manually entered into an Excel database. The photos and field notes total one terabyte

[1] "Update from ARCE: Current Research, Excavation and Conservation Projects in Egypt," *NILE* Magazine, October-November 2018, 59.
[2] "Update from ARCE," *NILE* Magazine, 59.

of data. Dr. Cooney is no longer collecting data. At this stage of her research she is interested in cleaning, curating, and sharing her data via publications and her research and educational website called WikiArtifact.

At the start of this quarter, we met with Dr. Cooney and her research team to review our data management plan from the winter quarter to determine next steps for data management and WikiArtifact. Dr. Cooney expressed concern about the financial sustainability of WikiArtifact, and wanted to ensure WikiArtifact can be financed and maintained for the long term. Last quarter, Dr. Cooney envisioned WikiArtifact as a visual, interactive online database of her research data, complete with visual tags that users could click on to view and search multiple data elements of the coffins, such as location, mythical imagery, and coffin materials. Dr. Cooney wanted WikiArtifact to be accessible, usable, and dynamic, with 3D imaging of her coffins and easy sharing via social media. The database would be collaborative, with vetted researchers adding their coffin-related data.

Accordingly, we recommended Omeka as a platform that could accommodate all of these functionalities. The platform is available via two routes: Omeka.org and Omeka.net. Omeka.org is a free, open-source platform, but requires the user to have their own server space.[3] Maintaining a server can be complicated and costly, so we investigated Omeka.net. With Omeka.net, users pay for server space hosted by Omeka based on their storage needs.[4] Omeka's support team gave us a quote of $5000 a year based on Dr. Cooney's need to store one terabyte of data. Even if Dr. Cooney were able to eliminate duplicate photos and reduce her storage needs to half as much (500 gigabytes), Omeka would cost $3000 a year.[5] Though Omeka checked several boxes off of

[3] "Omeka Classic," https://omeka.org/classic/. (Accessed June 6, 2019).
[4] "Pricing," Omeka.Net, https://www.omeka.net/signup. (Accessed June 6, 2019).
[5] "Omeka.Net Price List," 2018, Corporation for Digital Scholarship.

Dr. Cooney's must-have list, it proved to be too costly and the desired qualities of WikiArtifact needed to be reevaluated.

Accordingly, instead of envisioning WikiArtifact as a visual database, Dr. Cooney expressed interest in sharing her data through photo essays.[6] Dr. Cooney still envisions the project as being collaborative, and would like it to be shareable through social media. We took these concerns into consideration as part of this project, and looked into how to facilitate more cost-effective WordPress and Drupal sites for WikiArtifact (see the section entitled "Objective 2: Identify Server and Platform Options for WikiArtifact," starting on page 15).

II. Project Objectives

Through our conversations with Dr. Cooney and her research team, we identified two objectives for this project that will best prepare Dr. Cooney's data for the creation of WikiArtifact. First, we will clean the data in Dr. Cooney's Excel database. Currently, data is difficult to extract from the spreadsheet, in large part due to an inconsistent and unwieldy metadata schema as well as a lack of controlled vocabulary. A more strategic, consistent approach to metadata schemata and controlled vocabularies will correct these issues, and get the data in a more contextualized, standardized state. This will both aid in current use of the spreadsheet and it will prepare the data for safe upload into WikiArtifact.

Our second objective is to identify server options and platforms for WikiArtifact. Dr. Cooney and her team were very interested in understanding the technical options for WikiArtifact. Our research and recommendations for server and platform options will prioritize

[6] Kara Cooney, Interview with Dr. Kara Cooney and her research team, In-person, April 18, 2019.

the long-term stability of the data, funding limitations, as well as Dr. Cooney's vision for the project.

III. Objective 1: Cleaning the Excel database

The current Excel database houses all of the qualitative data Dr. Cooney has collected on 300 coffins throughout the world. While robust and largely functional, Dr. Cooney's research team has encountered problems when trying to extract data from the spreadsheet. We have identified two elements of the spreadsheet—the metadata schema and the controlled vocabulary—that could be expanded upon and standardized to streamline the spreadsheet, making it more navigable and usable.

A. Metadata Schema

The Excel database currently uses a homegrown metadata schema, tailored to the specific needs of Dr. Cooney's research. The schema is comprised of 15 fields, including information on the holding institution of the coffin (city, museum, accession number), descriptive metadata about the coffin (coffin type, coffin part, dating, provenance, Niwinski number, name(s) of the deceased, title), reuse information about the coffin (date examined, reuse score, type(s) of reuse), and miscellaneous notes (notes and other notes). These fields set a strong foundation for us to work off of. Ultimately, we created a revised schema that more accurately captures the complexity of Dr. Cooney's data, as well as created a crosswalk to Dublin Core to enable potential data sharing on a larger scale.

a. Revised metadata schema

We identified three issues in Dr. Cooney's homegrown schema that we sought to address with a revised schema: inconsistency, a lack of specificity, and insufficient description. When reviewing the spreadsheet, we saw that metadata fields were used inconsistently in some cases. The "provenance" field, for example, sometimes included the museum the coffin was housed at, the name of the excavation team, and information about the buyer and seller of the coffin.[7] This inconsistent entry is in part because the field of "provenance" is too broad. Many elements contribute to an item's provenance, include acquisition details, creation information, and the item's holding information. Broad terms ellide this specificity. This complicates data retrieval, as it is not clear what information will be found in each field. Further, during the initial stages of data collection, having more specific, granular metadata fields encourages more thorough and consistent data collection. In addition to the overly broad metadata fields that are inconsistently used, we found that Dr. Cooney's spreadsheet lacked some critical metadata fields, such as administrative data on access rights or provenance metadata for the data itself (not the coffin).

Ultimately, we wanted to create a metadata schema that was useful to Dr. Cooney, and streamlined features of her existing database. Our fully revised scheme can be found in the appendix on page 30. Our implementation of the schema on a set of data can be found in the appendix on page 37. As noted on both of these pages, the new scheme provides for provenance, descriptive, and administrative metadata. We have also defined each element in the revised metadata schema, so Dr. Cooney and her team can fully understand the new schema (appendix, page 30). Starred elements on this appendix item represent Dr. Cooney's initial elements that we

[7] Kara Cooney and Amber Wells, 2018, "Coffin List (FULL) 3.0."

carried forward to the revised schema. The other elements that are not starred, represent revisions and additions we made.

First, our revised scheme breaks apart several of Dr. Cooney's original fields in order to capture the nuance and complexity of the data. This granularity improves retrievability and entry consistency, and follows the principle of "atomizing" information—essentially, breaking up information so that each metadata field only contains one type of data.[8] An example of this can be found with the revisions we made to the "dating" field. Previously, Dr. Cooney's schema only had one field for describing the date of the coffins. This meant that the field was often filled with long, continuous blocks of texts, such as "19th dynasty to mid 20th dynasty ???"[9] In the revised schema, the "dating" field has been broken into five distinct fields: coffin start time period and coffin end time period, both accompanied by fields to add period descriptors (such as early, mid, late), as well as field to denote ambiguity about the period. From the above example of "19th dynasty to mid 20th dynasty ???", the information would be "atomized" into distinct chunks: 19th dynasty, mid, 20th dynasty, and ambiguous. Breaking "dating" into these five categories makes the Excel noticeably more navigable, as you can sort and filter for these facets by column.

Similarly, we broke out the "provenance" field to be more specific, adding fields such as "buyer," "seller," and "acquisition date." We also added a few categories for data provenance, including "data collector" and "reuse observations and explanation." These fields were missing from the previous spreadsheet, but are important for tracking the provenance of the data, thereby engendering trust in the data for external researchers who access WikiArtifact and wish to reuse the data.

---

[8] Carly Strasser, 2015, "Research Data Management," NISO Primer Series, Baltimore, MD: National Information Standards Organization, 5.
[9] Cooney and Wells, "Coffin List (FULL) 3.0."

The final two fields from Dr. Cooney's initial schema that we reworked were "notes" and "other notes." These sections functioned as catch-all fields within the database for miscellaneous information related to the coffins. We reviewed the content of these fields, and identified three commonalities: notes on file location, information on related coffins, and reuse observations for the coffin. We thus eliminated these two categories, which were vague and inconsistently utilized, and created these three new metadata categories. It is advised, however, that Dr. Cooney and her research team do their own review of the "notes" and "other notes" sections, to determine if there was any other categories information that should be added to the revised metadata schema. We are neither Egyptologists nor were we deeply entrenched in the data collection process. Dr. Cooney and her team may be able to identify additional patterns and commonalities within these two sections.

b. Dublin Core Crosswalk

In addition to revising Dr. Cooney's metadata schema to make the data easily retrievable and navigable, we also wanted to provide Dr. Cooney with the option to make her data more interoperable via a standardized schema. Last quarter, we suggested reviewing two potential standardized schemata for Dr. Cooney's data: MIDAS Heritage and Dublin Core. Developed by the Forum on Information Standards in Heritage and recommended by the Digital Curation Centre, MIDAS Heritage is a robust metadata standard for describing archaeological buildings, sites, and artifacts.[10] When reviewing the MIDAS Heritage standard, however, we found their categories and subcategories too detailed, and perhaps overwhelming to a research team with limited resources for information management.

[10] "MIDAS Heritage: The UK Historic Environment Data Standard," 2012, Forum on Information Standards in Heritage, https://historicengland.org.uk/images-books/publications/midas-heritage/midas-heritage-2012-v1_1/, 22 and 27.

Conversely, the Dublin Core metadata standard is a simple, low-cost metadata standard for digital objects. The schema was "designed to be extremely simple, flexible, and extensible" to encourage as wide adoption as possible. Dublin Core is comprised of just fifteen core elements, which are all optional and repeatable.[11] After evaluating its core elements, we determined that Dr. Cooney and her team would feel comfortable with the schema, especially as compared to other standardized schemata. Further, because the basic elements are simple and flexible, a wide variety of communities are more likely to use it—making the schema a good fit for Dr. Cooney's data, which spans the disciplines of art history, Egyptology, and archaeology.

Our Dublin Core crosswalk can be found in the appendix, on page 33. Should Dr. Cooney ever wish to make her data more shareable or interoperable, she now has a clear roadmap for doing so. Further, Dr. Cooney could encode these Dublin Core elements into the back-end of the WikiArtifact site—not the front-end—so that this metadata is searchable and retrievable without ruining the aesthetic or preferred term usage for metadata fields within the photo essays on WikiArtifact (for example, if Dr. Cooney would like to maintain the metadata field "types of reuse" instead of the more generic Dublin Core field "subject").

We encountered a few roadblocks when mapping Dr. Cooney's revised schema to Dublin Core. Overall, there were instances of unclear mapping. More specifically, we often found that we were unclear about whether we were describing the coffin, the photo of the coffin, or the dataset about the coffin. This problem arose in mapping to Dublin Core fields such as "contributor", "date," "type", and "language." Generally speaking, we prioritized describing the coffin, not the image of the coffin or the dataset. However, for "contributor," for example, we

[11] Stephen J. Miller, 2011, *Metadata for Digital Collections: A How-to-Do-It Manual*, London, UK: Facet Publishing, 51.

felt it was important to list Dr. Cooney as a "contributor" for her role in data collection, even though we were technically describing the coffin, not the data set about the coffin. While somewhat inconsistent, omitting this metadata would erase a lot of the context for the coffin within the setting of WikiArtifact. Similarly, the "type" of object per Dublin Core could be a physical object (the coffin itself) or a data set. The "language," too, could refer to the language on the coffins, or the language of the dataset. This ambiguity is a natural consequence of using Dublin Core, whose simplicity does not allow for nuanced description that could delineate these relationships. Nevertheless, the strengths of Dublin Core with regard to wide-scale adoption, simplicity, and interoperability make it the best fit for Dr. Cooney's data.

Additionally, some fields from Dr. Cooney's schema mapped onto several Dublin Core elements. For example, "name of the deceased" maps onto both "subject" and "description." Per Dublin Core's usage guideline, "subject" refers to "the topic of the content of the resource," often described in keywords or key phrases, while the "description" field is "an account of the content of the resource," serving as "a potentially rich source of indexable terms" that can use full sentences.[12] The "name of the deceased" sits between these two categories, without a clear-cut home.

It was also difficult to create the crosswalk without knowing the new context in which the data would be used. When crosswalking, it is not always necessary to map every element from the old schema to the new schema. It is only necessary to map the elements that are relevant to the new context. For example, the category "reuse score" would not always need to be mapped onto "description," if the new context in which the data is being used is not concerned with this

[12] "DCMI: Using Dublin Core," http://www.dublincore.org/specifications/dublin-core/usageguide/elements/. (Accessed April 27, 2019).

particular metric. In mapping the schema, we tried our best to be as inclusive as possible, to account for whatever new contexts the data may be used in.

There were also a few Dublin Core elements that we were not able to map to, as Dr. Cooney had not collected data on that front. "Format," for example, was unmappable, as Dr. Cooney did not collect data on dimensions of coffins or materials. Since Dublin Core does not require usage of each element, this is not a serious hindrance. But it does show how schemata are not always easily matched or aligned.

Finally, in completing the crosswalk we noticed that some categories from Dr. Cooney's schema do not exist in Dublin Core. This meant that some data collected is lost in the crosswalk to Dublin Core. For example, the location of the holding institution for the coffin, which in Dublin Core would be the location of the "publisher," did not make it during the mapping process. Fortunately, none of the affected elements were especially significant to understanding Dr. Cooney's research.

All of these issues are endemic to mapping and Dublin Core in general.[13] Despite this, the benefits of potential widespread sharing and interoperability make mapping a valuable exercise and potential option for Dr. Cooney's data.

B. Controlled Vocabulary

Dr. Cooney's Excel database suffers from inconsistent naming practices and data entry, in part due to turnover in research assistants and in part due to the nature of her data, which is subjective and thus can generate ambiguous descriptors. This has made it difficult to

[13] Mary S. Woodley, 2016, "Setting the Stage," In *Introduction to Metadata*, edited by Murtha Baca, http://www.getty.edu/publications/intrometadata/metadata-matters/.

systematically and efficiently extract information from the database, since each column had so many variant terms that filtering columns did not always retrieve accurate or complete results. Often when research assistants were attempting to create charts from the data, they had to manually review cells.

We took several steps to standardize vocabulary use within the spreadsheet. First, we collocated the terms. We ran the entire Excel spreadsheet through OpenRefine, a data cleaning application. This allowed us to collocate terms, and determine authority terms amongst variant terms. This provided a holistic view of what sort of terms were used and where the variation took place. We completed this for categories that required little Egyptology expertise, and have shared a list with Dr. Cooney and her research team to review and approve. However, there were quite a few categories that would require Egyptology expertise to understand the variant terms and their relationships. Accordingly, Dr. Cooney's research team will need to review these categories, and determine authority terms directly. Dr. Cooney has communicated that this will likely happen in Fall 2019, when they have more graduate student researcher support. This is a critical component of the data clean-up, and should be prioritized.

Our next step toward standardizing Dr. Cooney's vocabulary was to integrate the Getty Vocabularies, where applicable. Although not a perfect fit for her very specific data regarding coffin reuse, there are a number of fields within her spreadsheet that have been standardized via Getty Vocabularies. The Thesaurus of Geographic Names could be used for locations, while the Union List of Artist Names could be used for museum names. A full list of these recommendations, as well as other style and naming conventions, can be found in the appendix on page 35. Integrating the Getty Vocabularies into the spreadsheet will make Dr. Cooney's data

more interoperable; however, this is not as high of a priority as cleaning up the variant terms via OpenRefine or migrating the spreadsheet to the revised metadata schema, as the Getty Vocabulary standardization will only help the external researchers who are using these vocabularies. The other two data standardization tasks will help all users, including Dr. Cooney and her team. Accordingly, Dr. Cooney should only implement this recommendation if she has the time and resources to do so.

Finally, we found that our revised metadata schema solved some of our vocabulary issues, particularly with regard to ambiguity. The coffin dates, for example, are now broken into five columns that are granular enough to avoid variance. Where before, it was common to have variant dates like "early to mid 21st Dynasty," "early-mid 21st Dynasty," and "early/middle 21st Dynasty," now, the five metadata fields in the revised metadata schema encourage more consistent and standard inputs.[14]

C. Workflow and Implementation Recommendations

In order to jumpstart the database clean-up, we have revised a subset of Dr. Cooney's data with both the new metadata schema as well as our controlled vocabulary recommendations. These entries will also serve as an example of clean data, which Dr. Cooney and her team can review and refer to during their data cleaning process.

Dr. Cooney's team selected the coffins from Museo Egizio in Turin, Italy as the ideal starting point for WikiArtifact as they believe that Museo Egizio will be the most flexible in

[14] Cooney and Wells, "Coffin List (FULL) 3.0."

terms of image and data sharing.[15] Accordingly, we transformed the coffin entries from this museum, devising the following workflow for data-cleanup:

1) Migrate data from the old metadata schema to the revised schema.

a. Review and understand the new metadata schema. Read the definitions of the new metadata fields (page 30).
b. Understand the differences between the old schema and the revised schema.
c. Create a new Excel spreadsheet, with the new metadata fields as the header.
d. Copy and paste data from the old metadata spreadsheet to the new spreadsheet with the revised schema, moving data to the new metadata fields.

2) Standardize data through controlled vocabularies.

a. Import the spreadsheet into OpenRefine.
b. Correct for variant terms: select a column and then filter by text facet. On the left side panel, all variations within the column will appear. Cluster the terms, and then enter your preferred authority term. Merge and recluster the items.
c. Integrate the Getty vocabularies: review which columns should use the Getty vocabularies (page 35). Add these terms to the spreadsheet. For Getty terms that can be applied to multiple cells, you can apply them at scale in OpenRefine through clustering and assigning a preferred authority term again.

3) Clean the data. In addition to clustering variant terms, OpenRefine is a powerful tool for editing and transforming data.

[15] Cooney, Interview with Dr. Kara Cooney and her research team.

a. Eliminate white space. White space is extra spacing within a cell that is invisible to the eye, but can cause problems in data curation. Best practice for cleaning data is to eliminate white space.
b. Clean up the "types of reuse" field. In OpenRefine, you can break apart cells into multiple rows. This will allow for better manipulation, sorting, and filtering of this column, improving data retrieval. These edits do not export well into Excel, so this search functionality may be best done exclusively within OpenRefine.

We encountered a few difficulties when carrying out the data transformation for these 30 entries, which should be noted so that Dr. Cooney and her team can do their best to avoid them. During the initial steps of moving data from the old metadata schema to their new fields, we found the Excel spreadsheet to be somewhat cumbersome and not user friendly due to the sheer number of columns. We would recommend freezing the header so that team members do not need to scroll up to remember each column name. Column order could also be adjusted to best suit the team members' workflows. With regard to the controlled vocabularies, the Getty vocabularies rely on hierarchies to mark relationships. This may not be the most intuitive structure for new users. Fortunately, the two Getty vocabularies we are recommended—Union List of Artist Names (ULAN) and Thesaurus of Geographic Names (TGN)—are more straightforward on this front than a vocabulary like the Getty Art & Architecture Thesaurus, whose subject and topics are highly interconnected and more hierarchical than names and locations.

As for implementing the above workflow, we would highly recommend that Dr. Cooney and her team attend a training on OpenRefine. Utilizing OpenRefine as part of their data

clean-up will save Dr. Cooney and her research team an immense amount of time as the program is intuitive, powerful, and can transform data on a large scale. The Data Science Center at UCLA Library regularly hosts workshops on OpenRefine.[16] We would recommend attending one of these workshops or contacting the Data Science Center directly for training.

As described in the previous data management plan, data stewardship is an active, ongoing responsibility. The initial clean-up and transformation of data does not signal the end of data management practices. There are a few fields that are likely to change throughout the lifetime of the data, including "file location notes" and "access rights." Any changes to the data on these fronts should also happen to the Excel database. In fact, it is recommended that researchers revisit all data management documentation, including the previous plan and this report, on a weekly basis, to ensure follow-through and consistency as well as to record any updates.[17]

IV. Objective 2: Identify server options and platforms for WikiArtifact

A. UCLA IT Options

As addressed above, new server and platform options were needed in order to cut down on cost and fit the WikiArtifact's new format as photo essays. UCLA's IT department offered three options to build and host a website.

1. UCLA IT Option #1:

The most cost-effective option would be for Dr. Cooney to create her own WordPress site that could be coupled or uncoupled with a CPanel hosted by UCLA's IT department. A CPanel is

[16] "UCLA Library Events," UCLA Library, https://www.library.ucla.edu/events/data-cleaning-openrefine. (Accessed June 8, 2019).
[17] Strasser, "Research Data Management," 5.

a "web based hosting control panel provided by many hosting providers to website owners allowing them to manage their websites from a web based interface. This program gives users a graphical interface from which they can control their portion of the... server."[18] Hosting their own website on the CPanel means that Dr. Cooney's team would be responsible for maintaining the server; as such, they would be responsible for any security threats that UCLA deems concerning. This also means that Dr. Cooney and her team would need to use a platform like Drupal or WordPress for the site, and potentially pay a developer to add and maintain any plugins they want the site to have. Though this server option is the most cost effective at $28 a month, the responsibility of maintaining their CPanel could take up valuable resources like the time of graduate student researchers and funding for a developer to perform security maintenance.[19]

2. UCLA IT Option #2:

The second option UCLA IT offers is a service called Site Factory.[20] This option is $300 a month but comes with much more than server space.[21] Site Factory offers templates to create an interactive website, and UCLA's IT department used Site Factory to create several websites.[22] The style guide and templates use Drupal.[23] This option provides for a fast launch: through Site Factory, websites can be launched within 8 hours, versus 180 hours when creating your own custom site on cPanel. The downside of the Site Factory option is that it is not as customizable as

[18] "What Is CPanel? How to Use CPanel for WordPress Hosting," WPBeginner, https://www.wpbeginner.com/glossary/cpanel/. (Accessed June 11, 2019).
[19] Damon Wolf, Interview with Technical Account Manager | Information Technology Services at the University of California, Los Angeles, Phone, April 17, 2019.
[20] Ibid.
[21] Ibid.
[22] Ibid.
[23] Ibid.

Dr. Cooney may hope for. Additionally, this option is more costly per month and would require hiring a Drupal developer for roughly $110 an hour to develop the site and potentially maintain the site depending on the level of comfort Dr. Cooney and her team.[24]

3. UCLA IT Option #3:

The third option would be for Dr. Cooney and her team to manage their own server. To run a platform like Omeka with several plugins on a dedicated server, the cost would be nearly $200 per month for the server, and $0.09 for every GB of image storage (1TB would therefore be around $90/month).[25] This does not include the need for a web developer to create the site. This option is not recommended because it can be time-intensive, costly, and require some expertise.

B. Server and Platform Options Based on Dr. Cooney's New WikiArtifact Vision

In speaking with UCLA's IT department and Dr. Cooney, as well as through researching platforms, it became clear that the vision of WikiArtificat had to change based on the resources available to Dr. Cooney. All server and platform options had trade-offs such as functionality, budget, or sustainability. Dr. Cooney and her team made it clear that sustainability and usability are the top priorities.[26] After understanding the budget needed to support running a large amount of data on a server as well as customized APIs, Dr. Cooney and her team decided WikiArtifact should be a more streamlined site formatted as photo essays.[27] They also determined that WordPress is a good option to host WikiArtificat because Dr. Cooney's team has used and is comfortable with the platform.[28] Our team balanced the priorities of budget, accessibility, and

[24] Ibid.
[25] Damon Wolf, Interview with Technical Account Manager | Information Technology Services at the University of California. .
[26] Cooney, Interview with Dr. Kara Cooney and her research team.
[27] Ibid.
[28] Ibid.

sustainability, ultimately deciding that WikiArtifact should be built using WordPress, and Dr. Cooney should use WordPress server space to host the site.

WordPress offers a business option for $25 a month with unlimited storage space, which is necessary for Dr. Cooney's terabyte of photographs.[29] This pricing structure means a WordPress platform and hosting option would cost $300 year, a far cry from Omeka's $5000 a year. The pricing of WordPress combined with the team's general familiarity makes WordPress the most sustainable option for WikiArtifact. Furthermore, WordPress features that come with the business plan will be helpful to Dr. Cooney and her team, including social media compatibility, Google Analytics, live chat setup support, unlimited premium themes, and the option to install customized themes and plugins.[30] If Dr. Cooney does decide she wants to customize her WordPress site beyond the templates that WordPress offers, she may need to consider putting funding aside for a developer. Lastly, Dr. Cooney could start with a smaller and less costly WordPress site, and upgrade the site in the future when she uploads the full terabyte of data.[31]

V. Roadblocks and Future Recommendations

We have identified potential roadblocks on the data management side as well as with the technical implementation of WikiArtifact. Proactive and thoughtful stewardship will go a long way in addressing these issues.

A. Data Management

[29] "WordPress.Com Plans and Pricing – Get Started for Free Today!," *WordPress.Com* (blog), February 23, 2016, https://wordpress.com/pricing/.
[30] Ibid.
[31] Ibid.

a. Photo permissions

During our meeting with Dr. Cooney and her research team at the start of the quarter, Dr. Cooney communicated that museums can be protective around the dissemination of images and information about their holdings.[32] While Dr. Cooney took photos of the coffins personally and thus technically has copyright over that creative work, Dr. Cooney is very sensitive to museum needs and wants to maintain good working relationships with them. Accordingly, Dr. Cooney and her team will need to secure permissions from each museum in order to distribute coffin photos on a public website like WikiArtifact. Dr. Cooney could also watermark the photographs to prevent unauthorized dissemination, but securing photo permissions is the highest priority.

Dr. Cooney indicated that some museums may be more open to the idea of sharing the photos on WikiArtifact, such as Museo Egizio. Contacting such institutions should be prioritized in order to get WikiArtifact started. Once there is a critical mass of participants, other institutions may be more motivated to join and permit photo distribution.

In addition to recording these permission rights in the Excel database, it is best practice to also make it clear to external users how these photos can be used. Creative Commons and Rightsstatements.org both offer simple, standardized language regarding reuse rights that will help external users of WikiArtifact understand how they can use these coffin images.[33][34] Dr. Cooney could review these licenses and statements, determine which best fits for each rights situation, and then describe or directly link to the appropriate license or rights statement.

b. Codex

[32] Cooney, Interview with Dr. Kara Cooney and her research team.
[33] "About The Licenses," Creative Commons, https://creativecommons.org/licenses/. (Accessed June 6, 2019).
[34] "Rights Statements," https://rightsstatements.org/page/1.0/?language=en. (Accessed June 6, 2019).

Dr. Cooney's data is qualitative, as it consists of Dr. Cooney's observations and determinations regarding signs of coffin reuse. Further, her research covers a very niche subject, with terms like "decorative reuse" and various reuse scores not necessarily ubiquitously used or known. Dr. Cooney's research would benefit greatly from a codex, which defines every data entry within the sheet. With regard to dates, for example, a codex would define what "19th Dynasty" means, or "mid." With regard to types of reuse, a codex would define "name reuse" and "decorative reuse," for example. The codex should be made available on the WikiArtifact website, to help users of the data as well as potential contributors to WikiArtifact, better understand the data and ensure consistent usage of terms.

B. Technical Rollout of WikiArtifact

a. Staffing limitations

While Dr. Cooney's research team does have some background with WordPress, it is possible that Dr. Cooney may need to hire a WordPress developer to create WikiArtifact. WordPress has a strong user community and offers robust technical documentation for setting up and running WordPress sites; however, if Dr. Cooney wants to make best use of advanced plug-ins, beyond what WordPress templates offer, she may need to hire a software engineer.

b. Collaborative workflow controls

Dr. Cooney wants WikiArtifact to be a collaborative site where researchers can contribute their findings and data. Dr. Cooney and her team will need to vet researchers to confirm both the veracity and style of the data contributed. For instance, Dr. Cooney may wish to ensure that external data does not exhibit problems with term variance, and enforce authority terms for certain topics. Depending on the volume of data received and the resources at Dr.

Cooney's disposal, she may not be able to do this on the largest scale, but she could narrow vocabulary control to a preferred term for specific topics, such as dating or names. Dr. Cooney and her team should consider implementing a vetting form, such as Google Forms or the WordPress API Gravity Forms.[35] These forms allow site administrators to control content that goes onto WikiArtifact before it is published on the site.

c. Maintaining the WordPress site and its plug-ins

WordPress is an open source platform, which allows programmers to create plug-ins to integrate into WordPress sites.[36] The WordPress platform has regular updates to "fix bugs and ensure speed, security, and compatibility."[37] Similarly, many programmers who have created WordPress plug-ins will update those APIs when WordPress runs a site-wide update.[38] However, if Dr. Cooney utilizes plug-ins that are not frequently updated by the developers who created them, she may need to consider hiring a programmer to fix the code. This would be an added cost, and it should be a consideration when integrating plug-ins not offered by preset templates.

V. Project Timeline

Dr. Cooney's timeline is greatly dependent on graduate student researcher availability and funding. We have created two project timelines that Dr. Cooney can follow at whatever pace her staffing and resources allow. We suggest Dr. Cooney first implement the data management recommendations (in turquoise), and then begin creating WikiArtifact (in pink).

Phase 1: Data Management

---

[35] "Using the API Lead Form with Gravity Forms in Wordpress," Tripleseat Support, accessed June 7, 2019.
[36] "Why Is WordPress Free? What Are the Costs? What Is the Catch?," WPBeginner, January 22, 2019, https://www.wpbeginner.com/beginners-guide/why-is-wordpress-free-what-are-the-costs-what-is-the-catch/.
[37] "Why Is WordPress Free? What Are the Costs? What Is the Catch?," WPBeginner.
[38] Ibid.

A. Understand the Schema: In order to implement the new schema that we have created, Dr. Cooney and her team should review the schema and its definitions to better understand where information falls in the database.

B. Finalize the Controlled Vocabulary: There were many variant terms that we were not able to create authority terms for because we did not have the Egyptology expertise to do so. We have shared these lists of variant terms with Dr. Cooney and her team. They will need to review these and come up with a controlled vocabulary they feel comfortable with.

C. Generate Unique Identifiers: Creating unique identifiers for each coffins will greatly facilitate data retrieval and long-term stewardship. Currently, data regarding each coffin is located in the Excel, photo files, and WikiArtifact. Generating unique identifiers for each coffin—and integrating these unique IDs into all three of these spaces—will allow Dr. Cooney and her team to uniquely, efficiently, and unambiguously identify each coffin, no matter where they are digitally stored.[39] Linking photos, Excel entries, and WikiArtifact in this way will also be extremely helpful for new graduate student researchers who are not familiar with the coffin reuse database. Additionally, these unique IDs can be incorporated into a preferred citation for coffins within WikiArtifact, so that external users of the data can easily locate and access the correct coffin.[40]

[39] "On the Utility of Identification Schemes for Digital Earth Science Data: An Assessment and Recommendations | SpringerLink," accessed June 8, 2019, https://link.springer.com/article/10.1007%2Fs12145-011-0083-6.
[40] Ibid.

Unique identifiers can be created through the Online UUID Generator, a site that generates a unique identifier using a timestamp and the MAC address of the computer to make it truly unique.[41] Dr. Cooney and her team may want to create unique identifiers that have visible characteristics tied to the coffin; for example, each unique ID could begin with the initials of the city where the museum is located.

D. Cleaning the Data: The data within the database must be placed in the new schema using the controlled vocabulary once the previous steps are completed. When we tested out implementing the schema, we found that transforming ten coffin entries takes around 30 minutes. Therefore, 300 coffin entries will take approximately 15 hours to input.

E. Create a Codex: This step will also be very time consuming. It is not the most time-sensitive item, as it will be most helpful for external users of WikiArtifact, in understanding what the data means. It is, however, an important element for creating the necessary context needed for resharing data.

Phase 2: WikiArtifact Implementation

A. Contact Museums for Permissions: Without permissions from the museums, the vision of WikiArtifact as a public site is not viable. Accordingly, this is a very important first step in setting up WikiArtifact.

[41] "Online UUID Generator Tool," https://www.uuidgenerator.net/. (Accessed June 7, 2019).

B. Update the Database with Museum Permissions: Once the museums are contacted, the Excel should be updated with the museum's response or terms for sharing so that this information is in a centralized location. Keeping the museums' responses organized will also allow Dr. Cooney to understand if her current vision of sharing coffin photos publicly is achievable in the format she wants.

C. Create a WordPress Site: During this phase, Dr. Cooney can determine what plug-ins are needed and if she would like to hire a programmer to facilitate this process. A good starting point would be to eliminate duplicate photos of coffins, to streamline storage and identify which photos should be featured on the website. Duplicate photo sorting software exists for both Windows and Mac computers.[42]

D. Batch Upload Excel and Photos: Excel files can be batch uploaded to WordPress using a plug-in.[43] Batch uploading will prevent Dr. Cooney's team from manually inputting data. Similarly, WordPress offers a plug-in that batch uploads media files, including photos.[44]

E. Implement Vetting Forms: A form needs to be created that vets researchers who will contribute to WikiArtifact. Dr. Cooney can determine which metadata fields are necessary for each contributor to include, and embed these into the form. She can also determine which fields should comply with a controlled vocabulary, and either provide for that via a dropdown menu where applicable or send documentation to contributors with the authority terms. Such a controlled form will give Dr. Cooney and her team time

---

[42] "Photos Duplicate Cleaner on the Mac App Store," accessed June 8, 2019, https://itunes.apple.com/us/app/photos-duplicate-cleaner/id592704001?mt=12;%20https://www.ashisoft.com/blog/top-5-best-duplicate-photo-finder-to-delete-duplicate-photos/.

[43] "Import Spreadsheets from Microsoft Excel – WordPress Plugin | WordPress.Org," accessed June 9, 2019, https://wordpress.org/plugins/import-spreadsheets-from-microsoft-excel/.

[44] "How to Bulk Upload WordPress Media Files Using FTP," WPBeginner, January 10, 2018, https://www.wpbeginner.com/plugins/how-to-bulk-upload-wordpress-media-files-using-ftp/.

to confirm the accuracy of contributed data before it goes live on the site, as well as introduce some standardization via the schema elements and controlled vocabulary.

VII. Conclusion

Dr. Cooney's research on coffin reuse in Ancient Egypt provides a fascinating look into the economies, social conditions, artistry, and spiritual beliefs of Ancient Egypt. Dr. Cooney's research has revealed some stunning discoveries—for instance, coffin reuse rates during the 21st averaged 60%, suggesting that the practice was socially acceptable and legal. Some coffins even reveal multiple reuses.[45] After nearly a decade of traveling the globe, Dr. Cooney's findings offer a rare and invaluable look at Ancient Egyptian society, culture, and beliefs.

Active data management will help ensure the usefulness and preservation of all of the valuable data Dr. Cooney has collected. In addition to facilitating preservation, the recommendations within this report will help foster structured data sharing and peer-to-peer collaboration. The new schema and controlled vocabulary are crucial steps to increase uniformity and consistency in the data management process.

The collaborative, visual nature of WikiArtifact has "the potential to revolutionize how we approach object studies in archaeology, art history, and Egyptology."[46] The use of WordPress will make WikiArtifact a space for collaboration, centralized data sharing, and equitable access to cultural heritage materials, while allowing Dr. Cooney to stay within her budget. Furthermore, Dr. Cooney's team's familiarity with WordPress will play an important role in the launch and longevity of the project. Ultimately, conscientious and proactive data management is critical to facilitating WikiArtifact's goals and long-term success.

---

[45] "Update from ARCE: Current Research, Excavation and Conservation Projects in Egypt," *NILE* Magazine, 60.
[46] Cooney, Wells, and Campbell, 2018, "National Geographic: Storytelling and Technology," 2.

VIII. References

"About The Licenses." n.d. Creative Commons. Accessed June 6, 2019. https://creativecommons.org/licenses/.

Borgman, Christine L. *Big Data, Little Data, No Data: Scholarship in the Networked World*. Cambridge: MIT Press, 2015.

Cooney, Kara. Interview with Dr. Kara Cooney and her research team. In-person, April 18, 2019.

Cooney, Kara, and Amber Wells. "Coffin List (FULL) 3.0," 2018.

Cooney, Kara, Amber Wells, and Rose Campbell. "National Geographic: Storytelling and Technology," 2018.

"DCMI: Dublin Core Metadata Element Set, Version 1.1: Reference Description." Accessed April 27, 2019. http://www.dublincore.org/specifications/dublin-core/dces/.

"DCMI Type Vocabulary." n.d. Dublin Core Metadata Initiative. Accessed June 7, 2019. http://www.dublincore.org/specifications/dublin-core/dcmi-type-vocabulary/.

"DCMI: Using Dublin Core." n.d. Accessed April 27, 2019. http://www.dublincore.org/specifications/dublin-core/usageguide/elements/.

"Getty Thesaurus of Geographic Names." n.d. Getty Research Institute. Accessed June 8, 2019. https://www.getty.edu/research/tools/vocabularies/tgn/.

"Getty Union List of Artist Names." n.d. Getty Research Institute. Accessed June 8, 2019. https://www.getty.edu/research/tools/vocabularies/ulan/.

"How to Bulk Upload WordPress Media Files Using FTP." WPBeginner, January 11, 2018. https://www.wpbeginner.com/plugins/how-to-bulk-upload-wordpress-media-files-using-ftp/.

"Import Spreadsheets from Microsoft Excel – WordPress Plugin | WordPress.Org." Accessed June 11, 2019. https://wordpress.org/plugins/import-spreadsheets-from-microsoft-excel/.

"MIDAS Heritage: The UK Historic Environment Data Standard." Forum on Information Standards in Heritage, October 2012. https://historicengland.org.uk/images-books/publications/midas-heritage/midas-heritage-2012-v1_1/.

Miller, Stephen J. 2011. *Metadata for Digital Collections: A How-to-Do-It Manual*. London, UK: Facet Publishing.

Niwinski, Andrzej. 1988. *21st Dynasty coffins from Thebes: chronological and typological studies*. Mainz am Rhein: P. von Zabern.

"Omeka Classic." n.d. Accessed June 6, 2019. https://omeka.org/classic/.

"Omeka.Net Price List." 2018. Corporation for Digital Scholarship.

"On the Utility of Identification Schemes for Digital Earth Science Data: An Assessment and Recommendations | SpringerLink." Accessed June 9, 2019. https://link.springer.com/article/10.1007%2Fs12145-011-0083-6.

"Online UUID Generator Tool." n.d. Accessed June 7, 2019. https://www.uuidgenerator.net/.

"Photos Duplicate Cleaner on the Mac App Store." Accessed June 8, 2019. https://itunes.apple.com/us/app/photos-duplicate-cleaner/id592704001?mt=12;%20https://www.ashisoft.com/blog/top-5-best-duplicate-photo-finder-to-delete-duplicate-photos/.

"Pricing." n.d. Omeka.Net. Accessed June 6, 2019. https://www.omeka.net/signup.

"Rights Statements." n.d. Accessed June 6, 2019. https://rightsstatements.org/page/1.0/?language=en.

Strasser, Carly. 2015. "Research Data Management." NISO Primer Series. Baltimore, MD: National Information Standards Organization.

"UCLA Library Events." n.d. UCLA Library. Accessed June 8, 2019. https://www.library.ucla.edu/events/data-cleaning-openrefine.

"Update from ARCE: Current Research, Excavation and Conservation Projects in Egypt." *NILE Magazine*, November 2018.

"Using the API Lead Form with Gravity Forms in Wordpress." Tripleseat Support. Accessed June 9, 2019. http://tripleseat.zendesk.com/hc/en-us/articles/219006788-Using-the-API-Lead-Form-with-Gravity-Forms-in-Wordpress.

"What Is CPanel? How to Use CPanel for WordPress Hosting." WPBeginner. Accessed June 7, 2019. https://www.wpbeginner.com/glossary/cpanel/.

"Why Is WordPress Free? What Are the Costs? What Is the Catch?" WPBeginner, January 22, 2019. https://www.wpbeginner.com/beginners-guide/why-is-wordpress-free-what-are-the-costs-what-is-the-catch/.

Wolf, Damon. Interview with Technical Account Manager | Information Technology Services at the University of California, Los Angeles. Phone, April 17, 2019.

Woodley, Mary S. 2016. "Setting the Stage." In *Introduction to Metadata*, edited by Murtha Baca. http://www.getty.edu/publications/intrometadata/metadata-matters/.

"WordPress.Com Plans and Pricing – Get Started for Free Today!" *WordPress.Com* (blog), February 23, 2016. https://wordpress.com/pricing/.

IX. Appendix

A. Revised Metadata Schema Definitions

The following are definitions of each element of the revised metadata schema. Understanding these will facilitate consistent data entry. We have starred (*) elements that were carried forward from Dr. Cooney's initial metadata schema. Other elements that are not starred represent revisions and additions we made.

Provenance metadata

- **City of holding institution***: City in which the institution that holds the coffin is located. Best practice is to use the Getty Thesaurus of Geographic Names (TGN) vocabulary.[47]
- **Holding institution***: Institution that holds the coffin. Best practice is to use the Getty Union List of Artist Names (ULAN) vocabulary.[48]
- **Accession number***: The holding institution's unique identifier for the coffin.
- **Niwinski number***: Correlating coffin number in the Niwinski study.[49]
- **Acquisition date**: Date in which the coffin was acquired by the museum. Best practice is to use the format MM-DD-YYYY.
- **Purchase location**: Location in which the coffin was acquired by the museum. Best practice is to use the TGN vocabulary.
- **Seller**: Agent who sold the coffin to the holding institution. Best practice is to use the ULAN vocabulary.
- **Buyer**: Agent who acquired the coffin for the holding institution.. Best practice is to use the ULAN vocabulary.
- **Current collection**: Collection in which the coffin is currently housed in within the holding institution.
- **Excavation location**: Location in which the coffin was excavated. Best practice is to use the TGN vocabulary.
- **Excavation date**: Date(s) in which the coffin was excavated. Dr. Cooney can make a decision as to whether the date range or final date of excavation is preferred. Best practice is to use the format MM-DD-YYYY–MM-DD-YYYY.
- **Excavation team/agent**: Team or agent that excavated the coffin. Best practice is to use the ULAN vocabulary.
- **Date examined***: Date in which the coffin was examined for the purposes of this coffin reuse study. Best practice is to use the format MM-DD-YYYY.

---

[47] "Getty Thesaurus of Geographic Names," Getty Research Institute, https://www.getty.edu/research/tools/vocabularies/tgn/. (Accessed June 8, 2019).
[48] "Getty Union List of Artist Names," Getty Research Institute, https://www.getty.edu/research/tools/vocabularies/ulan/. (Accessed June 8, 2019).
[49] Andrzej Niwinski, 1988, *21st Dynasty coffins from Thebes: chronological and typological studies*, Mainz am Rhein: P. von Zabern.

- **Data collector**: Agent who collected qualitative data on the coffin as part of this coffin reuse study.
- **Reuse observations and explanation**: Observations about the coffin's reuse, and justification for why the coffin received its reuse score.

Descriptive metadata

- **Coffin type***: When we queried Dr. Cooney's team for this definition, we received the following: "if it's inner or outer of board or mask." This is an example of "definition by example," and is not best practice because defining something by listing examples of it does "not establish clear boundaries between what is and is not included in a concept."[50] We recommend that Dr. Cooney offer a more robust definition to facilitate better understanding of this metadata field.
- **Ambiguity and/or notes on Coffin Type**: Any additional notes on the coffin type that go beyond defining the actual coffin type.
- **Coffin part***: When we queried Dr. Cooney's team for this definition, we received the following: "is if it's only lid or case or a fragment." Again, this is "definition by example," and is not best practice. We recommend that Dr. Cooney offer a more robust definition to facilitate better understanding of this metadata field.
- **Ambiguity and/or notes on Coffin Part**: Any additional notes on the coffin type that go beyond defining the actual coffin type.
- **Coffin time period descriptor (early, mid, late)**: A general descriptor for the time period of the coffin as being either "early," "mid," or "late." Dr. Cooney and her team may wish to define these more specifically, either here or in the project's codex (by number of decades, for example, represented in each stage).
- **Coffin start time period**: The starting Dynasty in which the coffin could be dated.
- **Coffin end time period**: The ending Dynasty in which the coffin could be dated.
- **Coffin date ambiguity**: A field to denote any ambiguity or uncertainty about the dating. Use the term "Yes" to denote that the period is uncertain and "No" to denote that it is not.
- **Additional dates**: Any other dates associated with the coffin and its items. For example, if mummy linens placed on the body found inside the coffin are inscribed with dates that differ from the coffin date (ie. the date of the interment of the body differs from the date of the coffin).
- **Name(s) of the deceased***: Names of deceased person(s) who have used the coffin.
- **Title***: Where applicable, the coffins reflect the museum's title for the object. Sometimes the title is chosen by Dr. Cooney and her research team. An explanation here of when and why Dr. Cooney sometimes chooses a new title would be helpful.

[50] Christine L. Borgman, *Big Data, Little Data, No Data: Scholarship in the Networked World* (Cambridge: MIT Press, 2015), 19.

- **Reuse score***: A rating of Dr. Cooney's confidence in her ability to see coffin reuse on a scale from 0 to 3, with 3 being obvious and clearly visible evidence of reuse, 1 being only circumstantial, and 0 being no visible evidence of reuse. To clarify, a 0 score does not mean that a given coffin was not reused; it just means that Dr. Cooney cannot see evidence of that (evidence of reuse could be carefully covered by a carpenter, for example).
- **Type(s) of reuse***: Dr. Cooney's determination of the types of reuse in the coffin. Terminology should come from a controlled list of terms determined by Dr. Cooney.
- **Relation**: Any notes regarding the coffin's relation to other coffins on the list. Include the unique identifier of the other coffin in this field as well.

Administrative metadata

- **Unique identifier**: Unique identifier for each coffin generated by Dr. Cooney's team.
- **Rightsholder**: The institution or agent with which Dr. Cooney and her team correspond concerning access rights for the coffin photos and data.
- **Access rights**: Information from museum permissions concerning who is allowed to see coffin photos and data.
- **File location notes**: Notes about locations of photos and field note files, including when these files are missing.

B. Dublin Core Crosswalk

The following is a crosswalk that maps Dr. Cooney's revised metadata schema onto Dublin Core.

- DC: Title
    - KC: Title
- DC: Creator
    - Not listed on Dr. Cooney's current spreadsheet. It is recommended that this field be populated the "Unknown," as the creators of these coffins are not known.
- DC: Subject
    - KC: Types of reuse
    - KC: Name of the deceased
- DC: Description
    - KC: Coffin type
    - KC: Coffin part
    - KC: Name of the deceased
    - KC: Reuse score
    - KC: Acquisition date
    - KC: Purchase location
    - KC: Seller
    - KC: Buyer
    - KC: Current collection
    - KC: Excavation location
    - KC: Excavation date
    - KC: Excavation team/agent
- DC: Publisher
    - KC: Holding institution
- DC: Contributor
    - KC: Data collector
- DC: Date
    - KC: Coffin time period descriptor (early, mid, late)
    - KC: Coffin start time Period
    - KC: Coffin end time period
    - KC: Coffin date ambiguity
    - KC: Acquisition date
    - KC: Excavation date
    - KC: Date examined
- DC: Type

    - Not listed on Dr. Cooney's current spreadsheet. It is recommended that this field be populated with both PhysicalObject and Dataset. Both of these terms are taken from the recommended DCMI Type vocabulary.[51]
- DC: Identifier
    - KC: Unique identifier
- DC: Language
    - Not listed on Dr. Cooney's current spreadsheet. Where there is writing on the coffin, it is recommended that Egyptian is cited using the recommended standards for Dublin Core: RFC 3066 and ISO 39, which define primary language tags and subtags.[52]
- DC: Relation
    - KC: Relation
- DC: Coverage
    - KC: Coffin time period descriptor (early, mid, late)
    - KC: Coffin start time Period
    - KC: Coffin end time period
    - KC: Coffin date ambiguity
    - KC: Excavation location
- DC: Rights
    - KC: Access rights
    - KC: Rightsholder

[51] "DCMI Type Vocabulary," Dublin Core Metadata Initiative, http://www.dublincore.org/specifications/dublin-core/dcmi-type-vocabulary/. (Accessed June 7, 2019).
[52] "DCMI: Dublin Core Metadata Element Set, Version 1.1: Reference Description," http://www.dublincore.org/specifications/dublin-core/dces/. (Accessed April 27, 2019).

C. Controlled Vocabulary Guidelines & Examples
In addition to eliminating variant terms via OpenRefine, we implemented the following general guidelines for standardizing and controlling the vocabulary throughout the spreadsheet.

Style Conventions:

- Use sentence case throughout
- Use commas to denote multiple items (not plus signs or ampersands)
- For specific dates or date ranges in which the month, day, and/or year are known, use MM-DD-YYYY or MM-DD-YYYY–MM-DD-YYYY.
- For the "Coffin Date Ambiguity" field, enter either Yes or No
- For the "Coffin time period descriptor" field, enter either "early," "mid," "late," or "N/A"
- Do not leave any column blank. For data that was not recorded, use "N/R"

External Thesauri and Controlled Vocabularies:
The following thesauri should be used in the denoted fields where applicable. It is marked when use of these controlled vocabularies is required; otherwise, it is only considered best practice.

The Getty Thesaurus of Geographic Names:

- City of Holding Institution
- Purchase Location
- Excavation Location

The Getty Union List of Artist Names:

- Holding Institution
- Seller
- Buyer
- Excavation Team/Agent

When mapping elements onto Dublin Core, the following vocabularies should be used:

- DCMI Type Vocabulary:[53] Type (required)
- RFC 3066 and ISO 39:[54] Language

[53]"DCMI: Dublin Core Metadata Element Set, Version 1.1: Reference Description," http://www.dublincore.org/specifications/dublin-core/dces/. (Accessed April 27, 2019).
[54] "DCMI: Using Dublin Core," http://www.dublincore.org/specifications/dublin-core/usageguide/elements/.

The following is an example of controlled vocabulary work through OpenRefine, taking these stylistic guidelines into account.

The unbulleted terms are the authority terms that should be used in place of all of the variant terms, which are bulleted underneath that term. Items marked with an asterisk (*) denote data entries with extra details that would be placed in the newly created "Ambiguity and/or Notes on Coffin Type" column.

Dr. Cooney's team will need to review these recommendations, given their Egyptology expertise.

COFFIN TYPE

N/R

- Blank

Inner coffin

- Inner Coffin
- inner coffin
- Inner coffin (seems to be Stola)*

Inner coffin, mummy board

- Inner coffin + mummy board
- Inner coffin/Mummy board

Outer coffin

- Outer Coffin
- Outer coffin (?)*
- Outer(?) coffin*

Outer coffin, inner coffin

- Outer coffin + inner coffin

Outer coffin, inner coffin, mummy board

- Outer coffin + inner coffin + mummy board

Stola coffin, inner coffin

- Stola, Inner coffin

D. Example Set of Cleaned Data

PROVENANCE METADATA														
City of Holding Institution	Holding Institution	Accession Number	Niwinski Number	Acquisition Date	Purchase Location	Seller	Buyer	Current Collection	Excavation Location	Excavation Date	Excavation Team/Agent	Date Examined	Data Collector	Reuse Observations and Explanation
Turin, Italy	Museo Egizio di Torino	2212; CG 10112-10114	377	1822	N/R	B. Drovetti	N/R	N/R	N/R	N/R	N/R	06-07-2011	Kara Cooney	Decorative reuse. Very thin plaster layer. No traces of plaster for reuse. Very nice wood. One thick plank for mummy board. Coffin lid base is also one thick plank, but more boards used. Case bottom of coffin uses many pieces of local wood, but not important for viewer to see quality. Sides of case, though are thick, high quality wood. Different woods suggestive of reuse from older coffins. So, probably the plaster layer is so thin because the wood is so nice. This is a nice version of simple dot decoration, but no originality. Case sides very repetitive and generic. Name on the lid original in polychrome. Title also in polychrome.
Turin, Italy	Museo Egizio di Torino	2212; CG 10112-10114	377	1822	N/R	B. Drovetti	N/R	N/R	N/R	N/R	N/R	06-07-2011	Kara Cooney	N/R
Turin, Italy	Museo Egizio di Torino	2214; CG 10121	None	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R	Kara Cooney	N/R
Turin, Italy	Museo Egizio di Torino	2217; CG 10110	378	1822	N/R	B. Drovetti	N/R	N/R	N/R	N/R	N/R	06-07-2011	Kara Cooney	Just coffin case with no lid. Woman's coffin with no signs of reuse. Coffin wood insanely nice. Sides are each complete boards. Smaller pieces by head but just to create the round shape. No raised relief though. Thinner layer of plaster and no plaster underneath with any previous decoration. Nice would could be a marker that the coffin was reused from temple wood or other coffin wood. Similar in decoration to Hennatawy in Boston, but not raised relief. Most interesting things for defensive burial - no gilding in combination with really nice wood, probably imported. Probably orpiment and yellow ochre.
Turin, Italy	Museo Egizio di Torino	2219; CG 10117a-b	379	1822	N/R	B. Drovetti	N/R	N/R	N/R	N/R	N/R	06-07-2011	Kara Cooney	Decorative reuse, Blank space for name, gender modification. Breasts removed. Name space erased but part of name of title is still underneath. This is interesting and suggests reuse. Title is wsir nbt pr Smayt _____ r nb n ist mwt nTr. Stnage. Why erase part of the name? Is it part of a title? Then, why erase it? Nothing seems ever written...sparkly yellow bits in paint so maybe orpiment. But, there is yellow paint! And, on second glance, varnish! So, whole coffin has orpiment/yellow ochre. More ochre and thus mustard color to coffin. Not best craftsmanship, but blu pigment and orpiment (so type II coffin).
Turin, Italy	Museo Egizio di Torino	2221; CG 10116a-b	380	1822	N/R	B. Drovetti	N/R	N/R	N/R	N/R	N/R	06-07-2011	Kara Cooney	Decorative reuse, gender modification. Male coffin changed to woman's, beard hole covered. Usurpation only clear because of the beard hole under the chin! Cracking around the lappets of the wig suggests addition of plaster for headcloth. Text at feet ends with Wsir only so maybe they never added the name in. Raised plaster relief of sundisks and figures. Yellow paint. No breasts added to this usurpation by a woman of a man's coffin. With this restoration, if there was a 19th Dynasty coffin underneath, I wouldn't see it.
Turin, Italy	Museo Egizio di Torino	2222; CG 10111a-b	381	1822	N/R	B. Drovetti	N/R	N/R	N/R	N/R	N/R	06-07-2011	Kara Cooney	No lid to outer coffin. Not much interest in this piece as I remember.
Turin, Italy	Museo Egizio di Torino	2226; CG 10104a-b	382	1822	N/R	B. Drovetti	N/R	N/R	N/R	N/R	N/R	06-07-2011	Kara Cooney	The painting on the case side interior is unfinished! White line on red background. Masculine Issi on case right side! God with label of Isis. Pyramid with tree goddess is breast shaped while pyramid with Hathor of western mountain is sharp. This coffin is most interesting for male-female differences. For instance, strange that a male deceased is getting water from the tree goddess even though the coffin was made for the female Tabakenkhonsu. Female deceased is depicted once on each side and is carefully labeled.
Turin, Italy	Museo Egizio di Torino	2226; 10105	382	1822	N/R	B. Drovetti	N/R	N/R	N/R	N/R	N/R	06-07-2011	Kara Cooney	Mummy board has same style of raised relief (as 2226; CG 10104a-b inner coffin). If compare piece to piece, the inner coffin is higher in value than mummy board.
Turin, Italy	Museo Egizio di Torino	2227; CG 10115	383	1822	N/R	B. Drovetti	N/R	N/R	N/R	N/R	N/R	06-07-2011	Kara Cooney	Deceased in front of a number of dvinities. Nothing usurped is visble to me. Decent wood.
Turin, Italy	Museo Egizio di Torino	2228; CG 10119a-b	384	1822	N/R	B. Drovetti	N/R	N/R	N/R	N/R	N/R	06-06-2011	Kara Cooney	Multiple reuse, decorative reuse, markers of Ramesside, blank space for name. Whole surface of the lids is covered with a layer of linen and then plaster so no indication of white dress with red pleasts that must have been underneath. Dyn 19 according to feet and female shape of body. Label says "daily dress"... Sandals added to the feet in usurpation! The only instance of the dead's name is on the strip between the feet on the lid. Case sides--text space for naming female deceased is left out. Case right side--with tree goddess scene called nbt pr Smayt. In large format next to small caption we read Wsir nbt pr Smayt n imn but blank where name should be! Case left side in polychrome Wsir nbt pr Smayt n imn and erased name. But a mn sign is visible! So, maybe this is twice usurped.
Turin, Italy	Museo Egizio di Torino	2228; CG 10120	384	1822	N/R	B. Drovetti	N/R	N/R	N/R	N/R	N/R	06-06-2011	Kara Cooney	No text on mummy board at all!
Turin, Italy	Museo Egizio di Torino	2236-2237; CG 10101a-b	385	1822–1824	N/R	B. Drovetti	Museo Egizio di Torino	N/R	probably from TT 291	1818	B. Drovetti	06-05-2011	Kara Cooney	Decorative reuse. Painted plaster visible under surface on right shoulder; shoulders shaved down on inside to fit inner coffin, so not made as set. The traces of reuse are few, but one detail of painted plaster under coffin's right shoulder. Got good detail photographs. Plaster on lid much thicker than plaster on case. But, on both, the plaster is rough and lumpy, as if an older coffin was choseled down of its decoration, but didn't plane the wood down before they did the decoration.
Turin, Italy	Museo Egizio di Torino	2236-2237; CG 10102a-b	385	1822–1824	N/R	B. Drovetti	Museo Egizio di Torino	N/R	probably from TT 291	1818	B. Drovetti	06-05-2011	Kara Cooney	Decorative reuse. Coffin very small and shaved down on the inner shoulders, probably because the mummy didn't fit into the coffin with the lid. The inner coffin and outer coffin are way out of sync for a set. These can't have been made for one another but were pieced together. The inner and outer coffin also do not fit in terms of decoration. The outer has large sections of unvarnished white. The inner is completely varnished a dark amber. Outer coffin has the Nut - Geb scene in same place as teh inner coffin, but outer coffin has ithyphallic Geb.
Turin, Italy	Museo Egizio di Torino	2236-2237; CG 10103	385	1822–1824	N/R	B. Drovetti	Museo Egizio di Torino	N/R	probably from TT 291	1818	B. Drovetti	06-05-2011	Kara Cooney	N/R
Turin, Italy	Museo Egizio di Torino	2238; CG 10106a-b, 10107	386	1822	N/R	B. Drovetti	N/R	N/R	N/R	N/R	N/R	06-05-2011	Kara Cooney	Decorative reuse, name reuse. Thick smooth grained boards of wood. If not imported wood, then a finer native variety. Fair to say that 21st Dynasty wood is an improvement over most 19th Dynasty wood. The inside of the coffin is undecorated except for the floor, which has a goddess. That and the arms on the coffin lid suggest an early 21st Dynasty date, but N says mid-21st D. Nice example of simple dot version decoration. So, wood and decoration quality do not meet. Name of the case left side next to worshipping deceased added later in black paint under polychrome Wsir. Text in really bad shape and illegible to me, but they read Knsw-ms according to the label. But name of original deceased in written. Deceased on coffin case right side with arms upraised is called Ba-pw-n. Also, there is a small kneeling figure in the scale scene on case right with same name Ba-pw-n. Not in Ranke.
Turin, Italy	Museo Egizio di Torino	2238; CG 10106a-b, 10107	386	1822	N/R	B. Drovetti	N/R	N/R	N/R	N/R	N/R	06-05-2011	Kara Cooney	Decorative reuse, name reuse. Mummy board is one piece of wood like Hori. In fact, decoration is similar, too. Maybe related? Board has name reinscribed in black, too. Top starts with Dd mde in Wsir Bapwn Dd.f ... and then invocation to Nut for wings and imperishable ones.
Turin, Italy	Museo Egizio di Torino	CG 10109	387	N/R	N/R	N/R	N/R	N/R	Dayr al-Madinah	1905	Schiaparelli	06-08-2011	Kara Cooney	N/R
Turin, Italy	Museo Egizio di Torino	Suppl. 7715; CG 10108a-b	388	N/R	N/R	N/R	N/R	N/R	Dayr al-Madinah	1905	Schiaparelli	06-08-2011	Kara Cooney	Decorative reuse, gender modification. Man's coffin usurped by a woman. Visible because of the beard hole. Yellow paint. Varnish not evident. Male deceased depicted but not female. Usurpation of head cloth very interesting. Hard to see what they did but the face is poorly formed. Was it re-carved? Interior of coffin case is monochrome. Only red, black, and yellow, but underside of the lid has image of Osiris in polychrome. Really nice. Four sons of Horus as canopic jars. Pomegranates under offering table of one. Gorgeous solar scene of arms and tree and ba on case's left side. Very interesting coffin.
Turin, Italy	Museo Egizio di Torino	2226	383	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R	N/R	Kara Cooney	N/R

The following fields need to be standardized by Dr. Cooney's team, given their Egyptology expertise: title, name(s) of the deceased, types of reuse.
The following fields willbe completed by Dr. Cooney's team once they implement our plan: rightsholder, access rights, unique identifier.

DESCRIPTIVE METADATA									
Coffin Type	Ambiguity and/or notes on Coffin Type	Coffin Part	Ambiguity and/or notes on Coffin Part	Coffin Time Period descriptor (early, mid, late)	Coffin Start Time Period	Coffin Time Period descriptor (early, mid, late)	Coffin End Time Period	Coffin Date Ambiguity	Additional Dates
Inner coffin	N/R	Case, lid	N/R	mid	21st Dynasty	mid	21st Dynasty	No	N/R
Mummy board	N/R	N/R	N/R	mid	21st Dynasty	mid	21st Dynasty	No	N/R
Coffin	N/R	N/R	N/R	N/R	N/R	N/R	N/R	Yes	N/R
Outer coffin	N/R	Case	N/R	mid	21st Dynasty	mid	21st Dynasty	No	N/R
Inner coffin	N/R	Lid	This piece also has a mummyboard according to Notability file. The mummy board has a Blank Space for Name.	late	21st Dynasty	late	21st Dynasty	No	N/R
Inner coffin	N/R	Case, lid	N/R	late	21st Dynasty	late	21st Dynasty	No	N/R
Inner coffin	N/R	Case, lid	N/R	mid	21st Dynasty	late	21st Dynasty	No	N/R
Inner coffin	N/R	Case, lid	N/R	mid	21st Dynasty	mid	21st Dynasty	No	N/R
Mummy board	N/R	N/R	N/R	mid	21st Dynasty	mid	21st Dynasty	No	N/R
Outer coffin	N/R	Case	N/R	mid	21st Dynasty	late	21st Dynasty	No	N/R
Inner coffin	N/R	Case, lid	N/R	late	21st Dynasty	late	21st Dynasty	No	N/R
Mummy board	N/R	N/R	N/R	late	21st Dynasty	late	21st Dynasty	No	N/R
Outer coffin	N/R	Case, lid	N/R	mid	21st Dynasty	mid	21st Dynasty	No	N/R
Inner coffin	N/R	N/R	N/R	mid	21st Dynasty	mid	21st Dynasty	No	N/R
Mummy board	N/R	N/R	N/R	mid	21st Dynasty	mid	21st Dynasty	No	N/R
Inner coffin	N/R	Case, lid	N/R	mid	21st Dynasty	mid	21st Dynasty	No	N/R
Mummy board	N/R	N/R	N/R	mid	21st Dynasty	mid	21st Dynasty	No	N/R
Mummy board	N/R	N/R	N/R	mid	21st Dynasty	mid	21st Dynasty	No	N/R
Inner coffin	N/R	N/R	N/R	mid	21st Dynasty	mid	21st Dynasty	No	N/R
Outer coffin	N/R	Case	N/R	N/R	N/R	N/R	N/R	Yes	N/R

					ADMINISTRATIVE METADATA			
Name(s) of the Deceased	Title	Reuse Score	Type(s) of Reuse	Relation	Unique Identifier	Rightsholder	Access Rights	File Location Notes
N/R	N/R	1	Decorative Reuse; Mismatched Ledges;Mismatched Construction & Decoration	Coffin #277	N/R	N/R	N/R	N/R
N/R	N/R	1	N/R	Coffin #277	N/R	N/R	N/R	N/R
Anonymous	N/R	TBD	N/R	N/R	N/R	N/R	N/R	Notability file contains no notes.
Nesykhonsu	N/R	1	Mismatched Construction & Decoration	N/R	N/R	N/R	N/R	N/R
Anonymous	anonymous; Wsir nbt pr Smayt ___ ___f nb ist mwt nTr (?)	2	Decorative Reuse; Blank Space for Name; Gender Modification; Blank Space for Title; Plaster Modification	Coffin #379	N/R	N/R	N/R	N/R
anonymous man usurped by a woman	anonymous man usurped by a woman	2	Decorative Reuse; Gender Modification; Plaster Modification	Coffin #380	N/R	N/R	N/R	N/R
Bakenkhonsu	N/R	3	Plaster Modification; Wood Modification; Mismatched Ledges	Belongs in a set with Turin 2227 (Niwinski 383). Coffin #382	N/R	N/R	N/R	N/R
Tabakenkhonsu	N/R	0	No visible reuse	Coffin #382	N/R	N/R	N/R	N/R
Tabakenkhonsu	N/R	0	No visible reuse	Coffin #382	N/R	N/R	N/R	N/R
Tabakenkhonsu	N/R	0	No visible reuse	Coffin #381 A little confused with this. C.2226 is the inner and outer coffin. A few lines down, you have what appears to be another line for the same outer coffin.	N/R	N/R	N/R	N/R
chantress of Amun	N/R	3	Multiple Reuse; Decorative Reuse; Markers of Ramesside; Blank Spack for Name	Coffin #384	N/R	N/R	N/R	N/R
chantress of Amun	N/R	0	No visible reuse	Coffin #384	N/R	N/R	N/R	N/R
Butehamun	N/R	3	Decorative Reuse; Wood Modification; Plaster Modification; Mismatched Lid & Case; Mismatched Ledges	Coffin #385	N/R	N/R	N/R	N/R
Butehamun	N/R	2	Decorative Reuse; Wood Modification; Plaster Modification; Mismatched Lid & Case; Mismatched Ledges; Contextual Reuse	Coffin #385 This is another scenerior where the pices between inner, outer, and mummyboard are stylistically mismatched, and it is only provanance that tells us they go together. I don't know what type of reuse this is.	N/R	N/R	N/R	N/R
Butehamun	N/R	2	Markers of Ramesside	Coffin #385	N/R	N/R	N/R	N/R
Khonsumes, reused by Bapu	N/R	3	Decorative Reuse; Name Reuse; Mismatched Construction & Decoration; Mismatched Ledges	Coffin #386	N/R	N/R	N/R	N/R
Khonsumes, reused by Bapu	N/R	3	Decorative Reuse; Name Reuse	Coffin #386	N/R	N/R	N/R	N/R
Mwt-n-pr-imn	N/R	0	No visible reuse	Coffin #387	N/R	N/R	N/R	N/R
Herpeniset	N/R	1	Decorative Reuse; Gender Modification; Mismatched Leges	Coffin #388	N/R	N/R	N/R	N/R
N/R	N/R	0	Nothing usurped is visible	Coffin #383. Deceased in front of a number of divinities. Decent wood.	N/R	N/R	N/R	N/R