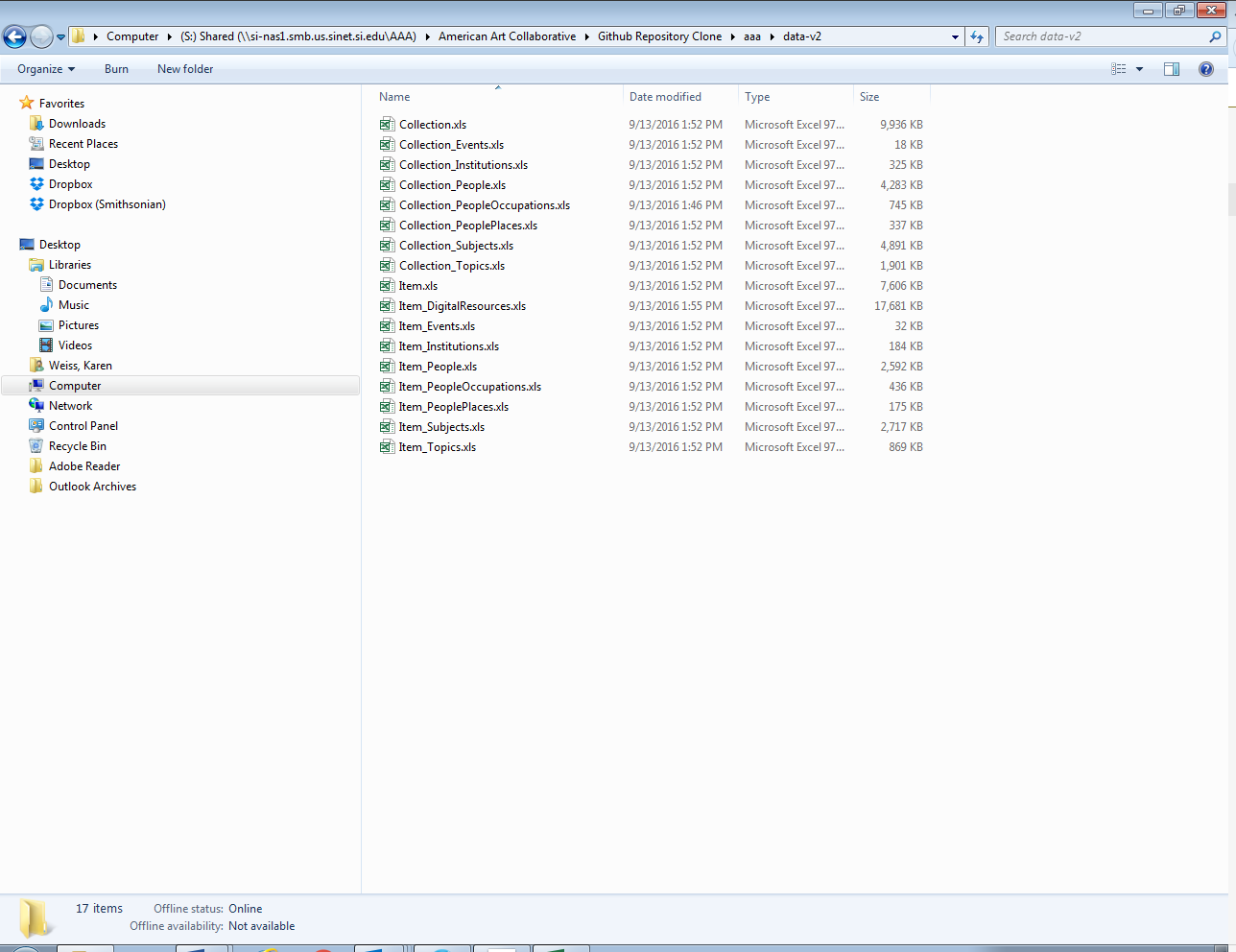
**Version 3 of AAA’s data contribution closely matches our version 2 format, and didn’t warrant a full rewrite of the documentation. This version of the document reflects minor changes, particularly field names. Version 3 data is stored in** <https://github.com/american-art/aaa/data-v3> .

The Archives of American Art’s Version 2 data contribution replaced the initial MARC XML and EAD XML initially submitted in April 2016. This version was loaded in a new GitHub Version 2 directory at   
<https://github.com/american-art/aaa/tree/master/version2/data>, and can now be found in <https://github.com/american-art/aaa/data-v2>. The dataset contains a total of 17 spreadsheets representing data for AAA’s collection level archival collections (Record Type=Collection); individually digitized documents (Record Type=Items).

Collection.xls comprises 8,280+ collection level descriptions. Item.xls comprises 15,560+ documents.

Collections and Items each have their own sets of associated spreadsheets to join the entities for People, Institutions and Events, and well as Topics. Items also has an additional spreadsheet for Digital Resources. People also have associated spreadsheets to join them with occupation(s) and their primary Place(s) of activity.

List of spreadsheets.



All Items have a relationship with a Collection (usually expressed as “forms part of” the collection) through a FK to the Collection ID in the Items.xls spreadsheet.

Collection.xls and Item.xls serve as the core descriptive information about a collection or item.

**COLLECTIONS (**8,280+ records)

The information on the Collections.xls spreadsheet includes:

* ID
* Medium – whether this is an archival record or part of AAA’s oral history program
* AAA Collection code (AAC may not use this but it’s included in case)
* FK ID for the Creator and in some cases, Co-Creator of the Collection. All collections have one Creator. It can be the name of a person, institution or event. The FK corresponds to the Collection\_People, Collection\_Institution or Collection\_Event spreadsheets which joins back to the Collection using the Collection ID.
* Collection Title
* Dates (collections typically have inclusive dates)
* Extent (collections are typically represented in linear feet) (sometimes Archives refers to this information as “Physical Description” or “Size”)
* Summary of the contents of the collection- Also known in archives as a “Scope and Content Note”
* BiographyBrief (this can also be a ‘historical’ note rather than biographical if the creator is an institution or event)
* BiographyFull: If there is an EAD finding aid for the collection, the spreadsheet includes the data from the EAD finding aid’s <biohistnote>tag (may include EAD tags).
* Provenance – narrative summarizing the background of the donation to AAA.
* Repository location (AAA’s address)
* URL to the persistent location on the www.aaa.si.edu website
* URL for a so-called “Representative Image” associated with the collection

**COLLECTION ENTITIES: SUBJECTS**

Entities that are *subjects* of the Collection are not represented as FK in the Collections.xls spreadsheet as creators and co-creators are. Instead, because there are often multiple entities, entities that are subjects are joined to the collection through the Collections\_Subjects.xls spreadsheet, and using the specific Collection ID and FK IDs for the entities in the Collection\_People, Collection Institution, Collection Event, and Collection\_Topic spreadsheets to join the data to the Collection as a subject.

Note, subject entities are derived from the Collection Summary and Biographical notes and are considered “index terms”. There is no representation for the specific relationship of the subject entities other than what can be deduced from reading the description as assembled from data in the Collections.xls spreadsheet.

**COLLECTION ENTITIES: PEOPLE – CREATORS AND SUBJECTS**

Collection\_People (that may be associated with a collection as a creator or subject) also expands to join with Collection\_PeopleOccupation and Collection\_PeoplePlace using the Person ID to provide information on the Person entities occupation(s) and primary place of activity (place names have been compiled from City, State and Country/region data for ease of use in this first pass).

**ITEMS**

The information on the Item.xls spreadsheet includes:

* ID
* FK to Collection ID of which the item is part
* Collection Title
* Display Date (items are typically single dates except for items like diaries, scrapbooks)
* Search Begin Date
* Search End Date
* General Format: AAA’s Designation of overall format. Mostly maps to AAT but AAA is currently working to separate Financial and Legal Records into separate formats.
* Specific Format: More specific document type – Generally maps to AAT
* Physical characteristic: Compiled data associated with the item relating to additional phys. Details
* Extent Number (ex. 1)
* Extent Type (ex. p.; or item; or v.)
* Item Size- h x w in cm.., rounded to nearest cm.
* Exact Duration – if General Format is Sound recording, videorecording or motion picture film
* FK Creator PersonID, or InstitutionID
* FK Co-Creator PersonID or Institution ID
* NOTE THAT THERE ARE NO CREATOR EVENTS OR CO-CREATOR EVENTS FOR ITEMS
* Description – Summary of the item. Not always present
* Citation - AAA’s preference for how an item should be cited
* Rights Statement – copyright and restriction information for the item
* Resource URL: Persistent URL to the item on the AAA website
* Current Location: AAA repository address
* Owner: AAA
* Representative Image (has representation): URL for the ‘representative’ image.
* Data datestamp

**ITEMS: SUBJECT ENTITIES**

Entities that are *subjects* of the Item are not represented as FK in the Item.xls spreadsheet as creators and co-creators are. Instead, because there are often multiple entities, entities that are subjects are joined to the collection through the Item\_Subjects.xls spreadsheet, and using the specific Item ID and FK IDs for the entities in the Item\_People, Item\_ Institution, Item\_Event, or Item\_Topic spreadsheets to join the data to the Item as a subject.

**ITEMS: PEOPLE ENTITIES**

Item\_People (that may be associated with an item as a creator or subjects) also expands to join with Item\_PeopleOccupation and Item\_PeoplePlace using the Person ID to provide information on the Person entities occupation(s) and primary place of activity (place names have been compiled from City, State and Country/region data for ease of use in this first pass).

**ITEMS-IMAGE URLS**

All items except for those of the General Format Sound recording, Videorecording or Motion picture film have URLS linked to the Item through the Item\_DigitalResources.xls spreadsheet using the Item ID. Note that if no order number appears in the spreadsheet, then the URLs for multiple images per item should be output in the order that they are listed in the spreadsheet for that item ID.