

# PhotoTech Materials Metadata

## Getty Research Institute

*Prepared by:*

Savannah Lake

December 2019

IS 498: Internship | Professor Blanchette

### **Contents:**

Project Summary	1
Documentation Overview	2
Data Content and Reconciliation Guidelines	3
Data Cleaning and Reconciliation Workflow	7
Data Cleaning and Reconciliation Workflow Diagram	12
GREL Cheat Sheet	13
Materials Reconciliation Terms and Formulas Spreadsheet	15

# Project Summary

## PhotoTech

The Photo Archive at the Getty Research Institute is an expansive collection, comprising of over two million photographic reproductions of artwork. Created in the 1970s, the archive was historically used by Getty staff and art historians to study artwork in the era before the internet. While accessing photographic reproductions of artwork is much easier now online, the items within the Photo Archive still hold strong research value, often containing provenance metadata. The PhotoTech project aims to bring the Photo Archive online, making both the images and their context more widely available for researchers.

The project faces some interesting challenges, largely due to the inconsistent processing and cataloging practices over the years. The collection as a whole has a finding aid, with varying levels of folder description. Approximately 14% of the collection has more robust item-level metadata recorded within a flat-file, command line database called the STAR database. The PhotoTech project also wishes to use computer vision to generate metadata from the stamps and handwritten notes on the photographs. The challenge will be reconciling these different sources of metadata to create an accessible and usable digital collection.

## Materials Data

My internship focused on the materials metadata within the STAR database, where catalogers described what materials, supports, and techniques were present in artworks. I was responsible for defining and documenting data content standards, and then cleaning metadata to those standards through OpenRefine. Once the metadata was cleaned, I was responsible for reconciling that metadata to the Getty Art and Architecture Thesaurus (AAT), in order to make the data more interoperable. This metadata will ultimately be transformed into the linked.art linked open data model. Throughout this process, I defined and documented a workflow. I also created a cheat sheet of General Refine Expression Language (GREL) syntax that I used to clean and reconcile the Photo Archive metadata. This documentation will help with the data cleanup effort within the PhotoTech project going forward, and could even benefit metadata cleanup efforts within the Getty Research Institute more broadly.

# Documentation Overview

## **Data Content Standards and Reconciliation Guidelines**

These guidelines document the decisions made during data cleaning and AAT reconciliation. The data content standards describe how the materials data should appear, and thus how the data was cleaned and standardized. This includes rules on syntax, punctuation, and capitalization, as well as how to format anomalous terms. The reconciliation guidelines identify and explain terms that we chose not to reconcile, as well as terms that have unique reconciliation target terms.

## **Data Cleaning and AAT Reconciliation Workflow: Outline and Diagram**

The workflow outline details the steps necessary for cleaning the materials metadata and reconciling it to AAT, providing explanations, goals, and sample GREL syntax. It also visually places the process within the broader PhotoTech Metadata Workflow diagram.

In addition to this written outline, the workflow diagram visualizes these data cleaning and reconciliation processes.

## **GREL Cheat Sheet**

The GREL Cheat Sheet lists all of the GREL syntax used to clean and reconcile the materials data. Additionally, the cheat sheet includes use cases relevant to PhotoTech data, as well as more thorough explanations to guide any new users of OpenRefine.

## **Materials Reconciliation Terms and Formulas Spreadsheet**

The reconciliation spreadsheet contains a list of all of the terms we are reconciling to (by material, support, and type) as well as all of the GREL formulas necessary to complete reconciliation. Any changes to these formulas can programmatically be built out there. Additionally, this spreadsheet contains all of the manual edits that will need to be executed during reconciliation.

# Data Content and Reconciliation Guidelines

## Contents

[Overview](#)

[Data Content Standards](#)

[General rules](#)

[Specific terms](#)

[Reconciliation Guidelines](#)

[Not reconciling \(too few entries, unlikely search points, or incompatible with the data model\)](#)

[Special reconciliation cases](#)

[To be determined](#)

## Overview

This document lists the data content standards for PhotoTech materials metadata, as well as decisions made with regard to reconciliation. These guidelines were drawn from historical documentation (namely the *Paintings Cataloging Worksheet Entry rules*, 1998), precedent in the data and other Getty Collections, and discussion with various members of the PhotoTech team, including Melissa Gill (metadata specialist and metadata lead for PhotoTech), Ann Harrison (special collections archivist with institutional knowledge of and experience with the STAR database), Ruth Cuadra (business applications manager and STAR database administrator), and Rob Sanderson (data architect and advisor for the linked open data model). When determining data content standards and reconciliation guidelines, the ultimate goal was to foster access and searchability while maintaining accurate description of the materials.

The data content standards apply to the first phase of the materials metadata project, in which we were cleaning legacy metadata from the STAR database. They describe how the materials data should appear in the free-text field, and thus how the data was cleaned and standardized. The data content standards include general rules on syntax, punctuation, and capitalization, as well as how to format anomalous terms.

The reconciliation guidelines apply to the second phase of the materials metadata project, in which we were reconciling terms to the Getty Art and Architecture Thesaurus. These guidelines identify and explain terms that we chose not to reconcile, as well as terms that have unique reconciliation target terms.

Below are working documents that outline some of these data decisions:

[STAR data cleaning weekly meeting agendas](#)

[Reconciliation check-in meeting](#)

[Paintings database proposed changes](#)

## Data Content Standards

### General rules

#### Capitalization

Always capitalize the first word in a field. Otherwise, do not capitalize unless the material is a proper noun.

#### Punctuation

Use commas only; no semicolons, periods, ampersands, or dashes.

Question marks should follow the end of a text string in parentheses.

##### **Oil on canvas (?)**

Oil on canvas?

Oil (?) on canvas

#### Term order

Term order should follow: medium, original support, any subsequent supports

Normalize term order when it does not change the semantics of the description. Use the most prominent occurrence if it aligns with the other content rules.

##### **Pen and bistre or ink, watercolor** (159 rows)

Pen and bistre or ink and watercolor (4 rows)

Watercolor, pen and bistre or ink (4 rows)

Pen and bistre or ink watercolor (2 rows)

For techniques, leave order as is if already incorporated semantically within the materials description.

Otherwise, if it is tacked on at the end in parenthesis, replace format with a comma.

oil on canvas (grisaille)

##### **Oil on canvas, grisaille**

Use either a comma OR an “on” for the support.

##### **Black chalk heightened with white on gray-brown paper**

##### **Black chalk heightened with white, light brown paper**

Black chalk heightened with white, on light brown paper

#### Multiple media values

Multiple media should be separated by commas. Do not use “and” to separate multiple media, unless it is a paired item in which the “and” connotes a relationship (eg. pen and ink, pen and wash, pen and bistre).

Pen, bistre, charcoal, and paint on paper

Pen and bistre and charcoal and paint on paper

##### **Pen and bistre, charcoal, paint on paper**

However, if there are only two media, separate them with an “and”

##### **Red chalk and black chalk heightened with white**

Red chalk, black chalk heightened with white

Repeat terms, if appropriate, in order to be exhaustive.

Black and red chalks on paper

##### **Black chalk and red chalk on paper**

Black, red chalk on paper

## Specific terms

Cradles: format information about cradles at the end of the description, with a comma.

**Oil on oak panel, no cradle**

Oil on oak panel (w/o cradle)

Oil on oak panel (no cradle)

Lined canvas: format information about lined canvases at the end of the description, with a comma.

**Oil on canvas, lined**

Oil on lined canvas

Oil on canvas (lined)

Oil on canvas lined

There are many instances of “heightened with white” — leave them as is. Although not likely, there is a possibility that the work is heightened with a material other than chalk, so it is best to not to be more explicit than the data indicates. Eg. Red chalk heightened with white, gray-blue paper

Maintain the “and” for the following terms: “pen and ink,” “pen and bistre,” and “pen and wash.”

Pen, brown ink, brown wash over black chalk

Pen and brown ink and brown wash over black chalk

**Pen and brown ink, brown wash over black chalk**

Trois crayons: Keep trois crayons, but also add red, white, and black chalks for materials reconciliation.

Trois crayons on beige paper

**Red chalk, white chalk, black chalk on beige paper, à trois crayons**

Terms with additional information in the materials field: Maintain materials information only. Move other notes to the following fields.

- Move to the “object notes” field:
  - Dates for when support materials were transferred; eg. Panel, transferred to canvas (1923)
  - If the artwork is detached; eg. Detached fresco, Fresco before being detached
  - Where the artwork is located; eg. Fresco, ceiling
  - States of the artwork; eg. Ruined fresco, Damaged fresco, Original panel
  - Corrections; eg. Chalk (called pencil)
- Move to “form” field:
  - Form of the artwork; eg. Fresco fragment, Tondo
- Move to “dimension notes” field:
  - Dimensions, especially if measurements are recorded; eg. panel (octagonal)
- Move to “marks and inscriptions” field:
  - Information about inscriptions; eg. signed and dated 1735
- Keep within the “materials field”:
  - Information that identifies multiple images; eg. Dark brown and black chalk (top) black chalk (bottom), Body color (37)
  - Clarifying information about the absence of a material; eg. Tempera on linen (no gesso)

Terms to leave as-is:

- Wood: Keep “wood” as is to maintain any context, but reconcile to panel.
- Gold ground: keep variation in statements as is, and reconcile materials to “gold ground”. Eg. Tempera on gold ground panel; Tempera on panel, gold ground; Gold ground on panel
- Pierre-blanche: leave as is, without reconciling to anything in AAT. Only four records.
- Painted sketch: Not necessarily a material, but keep as is.

## Reconciliation Guidelines

The emphasis for reconciliation was to reconcile with terms that would create meaningful access points in faceted searches. Accordingly, if a term would not be used often in faceted searches, or if the PhotoTech data did not have enough entries for a material to make it a meaningful faceted term, we did not reconcile the term. As all terms remain in the free-text field for materials, users will be able to locate these more obscure terms through keyword searching.

### Not reconciling (too few entries, unlikely search points, or incompatible with the data model)

- Colors
- Types of wood
- Cradles
- Lined vs unlined
- Gesso
- Pierre blanche
- Anything that is not a material or a technique. This includes items classified as instruments and implements in AAT, such as pens, reed pens, and lead point.
- Oil in “oiled paper” and “heightened with oil”
- Tracing paper should not be reconciled to tracing as a technique, because sometimes people use tracing paper without tracing on it.
- “Washed” should not be reconciled to “wash” as a material, because within this dataset it describes a technique used on the support (generally, “paper”), and is an unlikely search point.
- Plywood, given the timeframe of these pieces.
- Metallic netting
- Feather

### Special reconciliation cases

- When “gold” is mentioned as a material, it should be reconciled to “gold leaf.” Entries with “heightened with gold,” for example, still refer to gold leaf (not gold paint), making “gold leaf” the correct material.
- “Watercolor” and “gouache” have entries in the AAT in both the materials and technique hierarchies. They should be reconciled to their respective technique entries, as that is their more likely access points in user searches.
- To maintain both specificity for keyword searches and retrieval for faceted searches, use the following model for terms related to “panel”:
  - Wood: leave as “wood” in the materials field, reconcile to “panel”
  - Softwood: leave as “softwood” in the materials field, reconcile to “panel”
  - Board: leave as “board” in the materials field, reconcile to “panel”
- “Stone” should be reconciled to “stone (worked rock)” — Provenance Index did this as well.
- “Painted” should be reconciled to “painting (image-making)” technique as opposed to a specific type of paint material.
- For the handful of records described as “sepia” only, reconcile to “ink.”
- Silk should be reconciled to “silk (textile)” instead of “silk (general, animal material),” since it is unlikely artists would be working with raw silk.
- When materials are uncertain, such as “bistre or ink,” reconcile to both for best recall.

### To be determined

- Object reconciliation with [pen and ink drawings](#) and [pen and wash drawings](#)?
- “Gold ground” has been created in AAT within the materials hierarchy. If it moves to the techniques hierarchy, all instances of gold, gold leaf, and gold ground should have “gold leaf” as the material and “gold ground” as the technique.

# Data Cleaning and Reconciliation Workflow

## PHASE I: DATA CLEANING

While Photo Archive catalogers followed the *Paintings Cataloging Worksheet Entry rules* when cataloging in the STAR database, the materials data field is a free-text field, which means there is a substantial amount of variance in the data. The following steps outline how to standardize data in OpenRefine.

### 1. Load CSV into OpenRefine

When loading the data into OpenRefine, select UTF-8 as the character encoding to preserve diacritics. You will also need to update the column names, either in Excel or in OpenRefine, as data exports from the STAR database use placeholder text instead of the field name. Should you choose to update the headers in Excel, be sure to import the CSV correctly so that it maintains diacritics (Open new Excel workbook. Select Data > from text > delimited > comma > next tab > file origin: unicode UTF8 > next tab).

### 2. Execute GREL functions that apply to the entire dataset

Execute the following before clustering, as it will reduce the amount of discrepancies within the set:

- To sentence case (see [GREL Cheat Sheet](#))
- Trim white space
- Any overarching spelling or syntax choices as dictated by the [data content standards](#) and AAT (eg. “gray” vs “grey”; “and” vs “&”)

### 3. Cluster and clean dataset

This is where the bulk of the work will come in. Clustering will group similar variants of the same term together so that you can make mass edits for the correct term. Fingerprint and n-gram fingerprint clustering will likely yield the most results, but go through all keying function options as they are able to catch different typos and aspects of clustering. Follow the [data content standards](#) when cleaning data.

### 4. Calculate tail to determine cleaning goals

Depending on the dataset and capacity constraints, you may not be able to clean all of the data. Speak with the project manager to determine the cleaning goal. Once the bulk of the data has been cleaned, calculate the tail of the dataset to see how many records would need to be cleaned in order to reach whatever benchmark has been agreed upon with the project manager.



You can grab counts of the data from the text facet pane, and calculate the tail in Excel or Google Sheets. [Here](#) is the spreadsheet with calculations for the PhotoTech materials data as an example.

### 5. Spell check

Run a spell check on all of the materials statements in Google Sheets or Excel. You will catch errors like watercolor and yelllow that clustering and ngram did not catch in OpenRefine.

### 6. Spot check

This step provides another level of looking at the data to help catch any remaining errors.

### 7. Reingest clean data into STAR

Submit the cleaned data to the STAR database administrator for reingest.



## PHASE II: AAT RECONCILIATION

Now that the data is standardized and clean, you will need to reconcile it to AAT and prepare it for transformation into linked open data. To do this, you will need to get the data to a point in which individual terms can be split into different columns, as seen in this [target reconciliation template](#), in which each material, support, and technique in a materials statement is parsed out and identified by its AAT ID.

### 1. Create your reconciliation CSV

Your target reconciliation spreadsheet only needs to have one instance of each material statement, as it will live outside of STAR and not with the records themselves. Accordingly, instead of working on the entire dataset with over 6,000 instances of “oil on canvas,” all you need to have is a spreadsheet of all the unique materials statements, listing “oil on canvas” just once.

To create this, request an export from the STAR database with the now cleaned materials data. View this data in OpenRefine, and from the text facet of the materials field, grab all of the unique values (see step 4 of “Phase I: Data Cleaning” for a screenshot of how to do this). Paste these into a new CSV file.

In addition to these text-string statements, create a new blank column where you will do the reconciliation work. It is important that you prepare data for reconciliation in a separate column, so that you can refer to the free-text string while implementing the following transformations and confirm that the transformations are carrying forward correctly. Further, several of the transformations require filtering down the data through a text filter on the free-text string column.

Reload this CSV into OpenRefine. As ever, whenever loading data into OpenRefine, be sure to select UTF-8 as the character encoding to preserve diacritics.

### 2. Prepare text strings for parsing

The ultimate goal is to parse individual values out of the text string. To do this, make the text-string as standardized as possible, leaving out any unnecessary information. It should look like a list of items, separated by a common delimiter (in our case, a comma), that you could break out into several commas. To achieve this:

Make all text lowercase

    Edit cells > Common transformations > To lowercase

Remove all conjunctions and unnecessary text. This will vary by dataset, but could include with, and, on, transferred. A few examples below:

```
replace(value, " on ", ", ")
replace(value, " glued onto ", ", ")
replace(value, ", transferred to", ",")
replace(value, "transferred from ", "")
replace(value, ", and ", ", ")
replace(value, " and ", ", ")
replace(value, " heightening", "")
replace(value, " or ", ", ")
```

### 3. Remove irrelevant data

There is some data that, per our [reconciliation guidelines](#), we are not reconciling. There are also some terms that need to be removed in order to successfully apply the GREL syntax for parsing data. See the “manual edits” tab of the [Materials Reconciliation terms and formulas](#) spreadsheet for a full list of terms to remove, including when to do them (at this step, or at step 7 below).

To remove this data, you can use the GREL chomp or replace formulas, supplemented by manual edits when needed. See the [GREL Cheat Sheet](#) for example use cases.

### 4. Remove duplicates

Now, remove duplicates within cells. For example, a material statement of “black chalk, red chalk” needs to be reconciled to “chalk” just once.

First, add commas between each word:

```
value.replace(" ", ", ").replace(",,", ",")
```

Remove duplicates:

```
value.split(',').uniques().join(',')  
value.replace(",,", ",")
```

While the above was useful for removing duplicates, it did erroneously insert commas between two-word terms (eg. India ink, body color). You can correct for these errors with formulas listed in the “manual edits” tab of the [Materials reconciliation terms and formulas](#) spreadsheet. To bring up the appropriate entries, use a text filter on the “media\_mat 1” column and then conduct the transformation on the “clean\_media\_recon” column. For example, filter “gold leaf” in the “media\_mat 1” column, and then in “clean\_media\_recon” run Edit Cells > Transform > replace(value, "leaf", "gold leaf")

### 5. Assign values to material, support, and technique columns

Create columns for materials, supports, and techniques:

```
add columns based on current column, set to null, name: material_id, support_id, technique_id
```

In the material\_id column, run the materials GREL expression from the “material\_ID formula” tab of the [Materials reconciliation terms and formulas spreadsheet](#). It should look like:

```
if(cells['clean_media_recon'].value.contains("oil"),",oil paint (paint)", "") +  
if(cells['clean_media_recon'].value.contains("tempera"),",tempera", "") ...
```

In the support\_id and technique\_id columns, run their corresponding formulas from this Excel.

### 6. Remove duplicates

Some elements repeat (for example, wood and panel both transform to panel). Eliminate these with:

```
value.split(',').uniques().join(',')
```

### 7. Correct incorrectly reconciled data

Manually correct terms as designated in the “manual edits” tab of the [Materials reconciliation terms and formulas spreadsheet](#). To remove this data, you can use the GREL chomp or replace

formulas, supplemented by manual edits when needed. See the [GREL Cheat Sheet](#) for example use cases.

When making these edits, be sure to review the original materials statement (“clean\_media\_recon” and/or “media\_mat\_1”) to be sure you are not overcorrecting. For example, material statements with “cardboard” in them are incorrectly reconciled to both “cardboard” and “panel,” because the word “board” is within “cardboard,” and the board is reconciled to panel. You’ll thus need to remove “panel” from reconciled supports. However, there could be a materials statement that includes both panel and cardboard—for example, “Oil on cardboard, laid on panel.” These statements *should* be reconciled to both cardboard and panel.

## **8. Split material\_id, support\_id, and technique\_id by separator**

You will now have three columns, each with a list of the corresponding terms for that materials statement that should look something like: charcoal,chalk,pencil

In order to run the reconciliation plug in, split these columns apart so each term is in its own column:

Edit column > Split into several columns > By separator > *set to a comma*

## **9. Run the AAT reconciliation plugin**

The full documentation of how to run the plug-in is here:

<http://www.getty.edu/research/tools/vocabularies/obtain/openrefine.html>

Run the plug-in on each column: material\_id 1, material\_id 2, etc; support\_id 1, support\_id 2, etc; technique\_id 1, technique\_id 2, etc.

Complete all steps of the documentation, including choosing the correct AAT term and creating new columns with the AAT ID. When choosing correct AAT terms as part of the reconciliation, it can be helpful to run a text facet on the original statements to confirm that all terms were correctly reconciled.

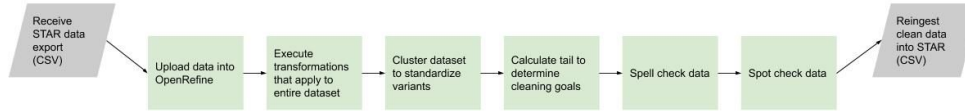
## **10. Create a clean AAT concordance table**

Once you have reconciled all the terms, format the CSV for use as a concordance table. For reconciliation, the linked data model only requires the information as laid out in the “Transform Template” tab of the [target reconciliation template](#).

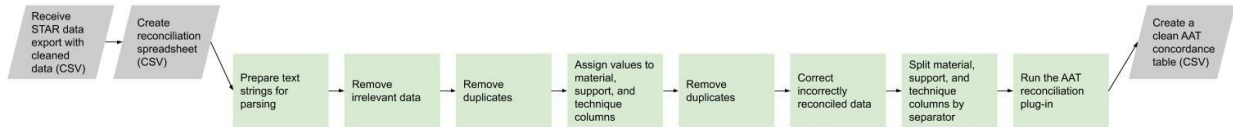
# RELATED MATERIALS

This entire workflow has been diagrammed [here](#).

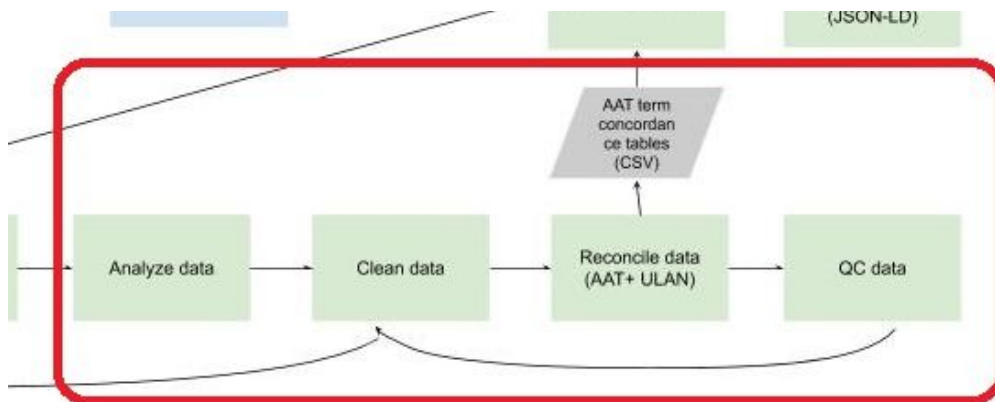
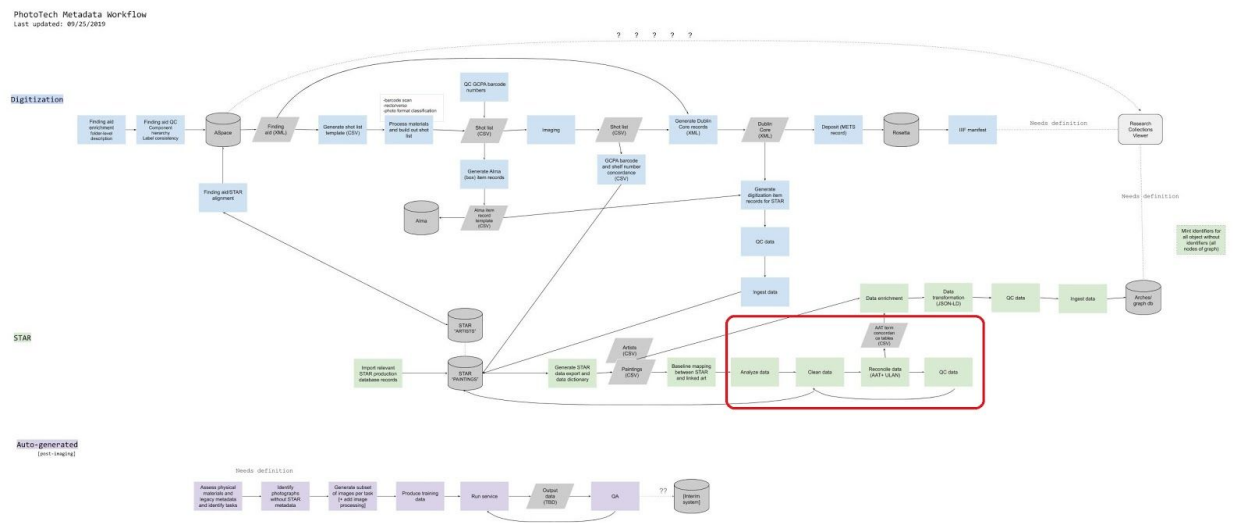
## Phase I: Data Cleaning



## Phase II: AAT Reconciliation

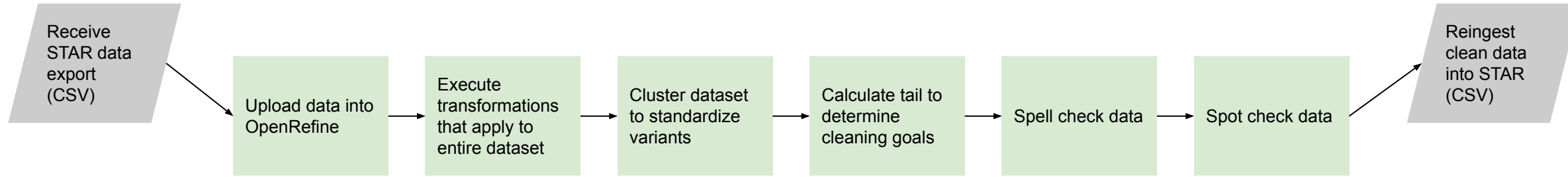


Within the broader [PhotoTech Metadata Workflow](#) diagram, these workflows fit in here:

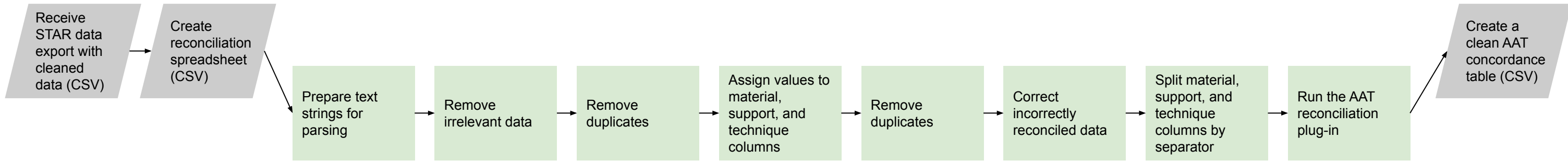


# PhotoTech Data Cleaning and Reconciliation Workflow | Materials Data

## Phase I: Data Cleaning



## Phase II: AAT Reconciliation



# GREL Cheat Sheet

## Phase I: Data Cleaning

### To sentence case

This function transforms text within a column to sentence case, the preferred format for PhotoTech data. This is helpful to do before clustering, to reduce clusters resulting from discrepancies in capitalization.

```
toUppercase(substring(value,0,1))+toLowercase(substring(value,1))
```

### Replace

This function replaces text within a column with desired text, replacing f with r:

```
replace(value, "string f", "string r") OR value.replace("string f", "string r")
```

This is helpful to do before clustering, to implement known standardization rules at the outset. It is also useful throughout data cleaning, when you identify trends in the data that need to be corrected.

For example, if you know that in AAT "gray" is preferred to "grey", you can use this function to standardize.

Text filter > grey

*[this step is optional, but will make transformations more apparent when typing the GREL formula, so you can ensure it is implementing properly]*

```
replace(value, "grey", "gray")
```

```
replace(value, "Grey", "Gray")
```

*Be sure to also catch capitalized words that begin strings.*

Further, this function can correct errors that clustering does not catch, such as ampersands.

Text filter > &

```
replace(value, "&", "and")
```

### Chomp

This function removes a specified string from the end of the string if said string is there:

```
chomp(value, "string") OR value.chomp("string")
```

Chomp can be an alternative to Replace when there's a possibility of replacing the incorrect part of the string, as the Chomp function localizes any changes to the end of the string.

For example, if you would want to standardize the following strings to all end with (?):

Oak and panel?

Red chalk and black chalk (?)

Canvas?

Using Replace [value.replace("?", " (?)")] would incorrectly alter the second string, which is already correct. Instead, use chomp and add the desired text string:

```
chomp(value, "?") + " (?)"
```

## Phase II: AAT Reconciliation

### Contains and If Control

In order to use the AAT Reconciliation OpenRefine plug-in, each material within a material statement must be parsed out into separate columns. This can be accomplished through a Contains function paired with an If control.

The Contains function checks if a specified string is present, returning true or false:

```
contains(value, "string") OR value.contains("string")
```

The If control evaluates whether an expression is true, implementing an expression if it is true and another expression if it is false:

```
if(expression o, expression eTrue, expression eFalse)
```

Within our use, the contains function checks a materials statement to see if it contains a certain value. If it does contain that value, it will populate a new column with that value. If it does not contain that value, it will populate a new column with nothing. For example, to check if the column named "clean\_media\_recon" contains the word "fresco", and to populate the current column you are running the transformation on with "fresco" if it does and nothing if it doesn't, use the following formula:

```
if(cells['clean_media_recon'].value.contains("fresco"),"fresco", "")
```

We have programmatically constructed this formula for the materials data in the [Materials reconciliation terms and formulas spreadsheet](#) spreadsheet.

### Remove duplicates

For any materials statement, we only need to reconcile to a term once. We can remove repeated materials with the following syntax. Since each material in our formatted statements is separated by a comma, the following splits the statement by commas, extracts the unique values, and rejoins them with a comma:

```
value.split(',').uniques().join(',')
```

### Removing irrelevant data

Part of reconciliation is removing items that could be incorrectly reconciled to an AAT term. Here, you can use the Chomp and Replace formulas again.

For example, to remove "tin" from strings like "watercolor,tin,oil" and "charcoal,tin", you can use:

```
chomp(value, ",tin")  
replace(value, ",tin,", ",")
```

### Resources:

VRA Conference's [GREL Expressions for visual resources management Common transformations](#) (including information on replace and contains)  
[GREL Controls](#) (including information on if statements)

<b>general term</b>	<b>AAT target term</b>	<b>notes</b>				
oil	oil paint (paint)					
bistre	bistre					
ink	ink					
charcoal	charcoal					
wash	wash					
chalk	chalk					
lead	graphite (mineral)	"black lead" captured here; when you split the terms to deduplicate, "black lead" becomes "black, lead"				
pencil	pencil (marking material)					
tempera	tempera					
graphite	graphite (mineral)					
india ink	india ink					
sinopia	sinopia					
stone	stone (worked rock)					
ground	ground					
gold	gold leaf					
gold leaf	gold leaf					
gold ground	gold ground					
pastel	pastel (material)					
<b>targeted term</b>	<b>AAT target term</b>	<b>notes</b>				
sepia	ink	Ignore if "sepia" precedes another material (wash, ink). If it is by itself, reconcile to ink.				





<b>term</b>	<b>AAT target term</b>	<b>notes</b>
panel	panel	
softwood	panel	
board	panel	
wood	panel	
cardboard	cardboard	
paper	paper	
canvas	canvas	
copper	copper	
metal	metal	
masonite	masonite	
vellum	vellum	
linen	linen	
slate	slate	
glass	glass	
tin	tin	
bronze	bronze	
cloth	cloth	
ivory	ivory	
leather	leather	
parchment	parchment	
alabaster	alabaster	
marble	marble	
burlap	burlap	
silk	silk (textile)	



<b>term</b>	<b>AAT target term</b>	<b>notes</b>
fresco	fresco	
silverpoint	silverpoint	
grisaille	grisaille	
brush	brush	
aquarelle	aquarelle	
etching	etching	
traced	traced	
metalpoint	metalpoint	
engraving	engraving	
varnished	varnished	
embroidering	embroidering	
à trois crayons	à trois crayons	
intarsia	intarsia	
woodcut	woodcut	
painted	painted	
gouache	gouache	
watercolor	watercolor	
body color	gouache	



Before splitting clean_media_recon into material, support, and technique columns <a href="#">(step 3 of the reconciliation workflow)</a>							
Remove from clean_media_recon							
any cradles and their materials							
lead point							
Additionally, you'll need to prevent overactive reconciliation—for example, reconciling "tin (metal)" to the term "tinted".							
The following terms have potential for this problem, and should be removed as designated (steps 3 and 7 of the reconciliation workflow).							
term	incorrect attribution	remove BEFORE splitting recon_data (step 3)	remove AFTER splitting recon_data (step 6)				
wood	plywood	plywood					
tin	tinted, painting, netting, writing, tint, acquaint	tinted, netting, writing, tint, acquaint	painting				
metal	metallic						
oil	oily, oiled, heightened with oil	oily, oiled, heightened with oil					
cardboard	panel (because board is reconciled to panel, and "board" is in "cardboard")		panel				
ink	pink, pinkish	pink, pinkish					
gold	golden, gold leaf when it should just be gold ground	golden	gold leaf when it should just be gold ground				
ground	background	background					
gold ground	ground		ground				
washed	wash	washed					
metalpoint	metal		metal				
Add back two-word terms after removing duplicates (step 4 of the reconciliation workflow)							
term	formula	notes					
black lead		no need -> captured in "lead", see "material_terms" in replace(value, "	lead	"	black lead	)	replace(value, "lead", "black lead")
body color	replace(value, "body", "body color")		body	"	body color	)	replace(value, "body", "body color")
india ink	replace(value, "ink", "india ink")		ink	"	india ink	)	replace(value, "ink", "india ink")
gold leaf	replace(value, "gold", "gold leaf")		leaf	"	gold leaf	)	replace(value, "leaf", "gold leaf")
gold ground	value.replace("ground.", ".") replace("gold", "gold ground")	DO NOT use this formula on entries of just "ground"	ground	"	gold ground	)	replace(value, "ground", "gold ground")
a trois crayons	replace(value, "crayons", "a trois crayons")		crayons	"	a trois crayons	)	replace(value, "crayons", "a trois crayons")
After splitting recon_data into material, support, and technique columns (step 7 of the reconciliation workflow)							
Remove colors and other terms that have been confused for materials or supports:							
ivory	"ivory paper" should not be reconciled to "ivory"						
gold	"gold pen" should not be reconciled to "gold leaf"						
copper	no confusion with copper in this data set, but others have the potential to conflate color with material here						
bronze	no confusion in bronze this data set, but others have the potential to conflate color with material here						
tracing paper	should not be reconciled as a technique						