

Marquette University

From the Selected Works of Rose Fortier

May 11, 2012

Staff Sourcing for Digitization

Rose Fortier, *Marquette University*



Available at: https://works.bepress.com/rose_fortier/13/

Staff Sourcing for Digitization

The background of the slide is a dark, abstract image featuring several bright, white light rays or beams that originate from the right side and fan out towards the left. These rays create a sense of depth and focus. Below the main image area, there is a solid, horizontal grey bar that spans the entire width of the slide.

What does it even mean?

- outsource – To send out (work, for example) to an outside provider or manufacturer in order to cut costs. (*American Heritage College Dictionary*, 4th ed.)
- staff source – To send out work to staff members outside of Digital Projects to increase productivity. (Rose Fortier, 1st ed.)

How it works @ MPL

- It works better for the digitization of some types materials than for others
- Digitization steps
 - Scanning
 - Image editing
 - Text creation (depending on the project)
 - Metadata creation
 - Import to content management software (we use CONTENTdm)

How it works @ MPL

- We staff source primarily for metadata creation
- Can also do it for text creation if not using Optical Character Recognition (OCR)
 - We use staff to proofread text created using OCR

Setting it up

- Need a (virtual) space to set up what staff will be working with
 - We use a common drive on the local network that is accessible by Neighborhood Library and Central Library staff
- Need the components that staff will work on
 - We have a few basic components:
 - Image
 - Text (if necessary for collection)
 - Database
 - Controlled vocabularies

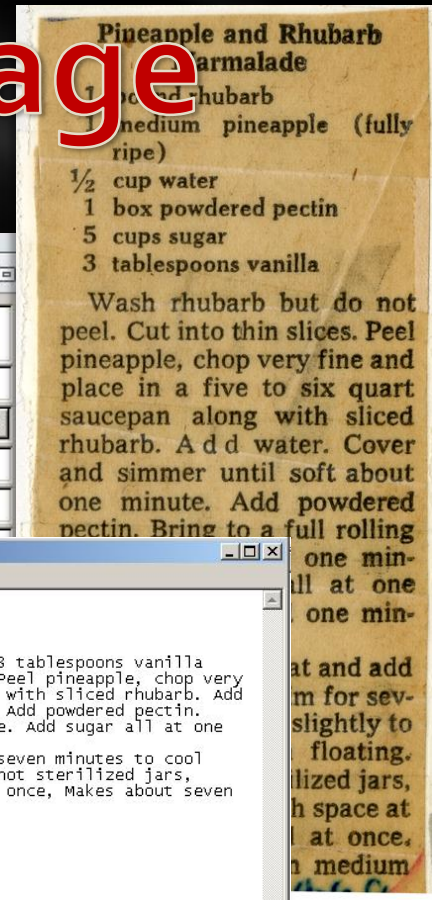
Image

Database

Historic Recipes

ID	1	Subject	Cooking, American; Baking; Pastry
Title	State Fair Cream Puffs	Preparation Time	
Meal	Snack	Newspaper	Milwaukee Journal
Course	Dessert	Date Appeared	1980-07-30
Ethnicity	American	Relation	Recipe_0004.jpg
Main Ingredient	Cream	Original	recipe_3461.txt - Notepad
Type of Dish		Type	Pineapple and Rhubarb Marmalade
Seasonal	Summer; State Fair	Has File	1 pound rhubarb 1 medium pineapple (fully ripe)
Recipe Source	Wisconsin State Fair	Collected	1/2 cup water 1 box powdered pectin 5 cups sugar 3 tablespoons vanilla
Transcript		Language	Wash rhubarb but do not peel. Cut into thin slices. Peel pineapple, chop very fine and place in a five to six quart saucepan along with sliced rhubarb. Add water. Cover and simmer until soft about one minute. Add powdered pectin. Bring to a full rolling boil and boil hard one minute. Add sugar all at one time and boil hard one minute. Remove from heat and add vanilla, stir and skim for seven minutes to cool slightly to prevent fruit from floating. Ladle into hot sterilized jars, leaving one-half inch space at top of jar s. Seal at once, Makes about seven medium glasses

Record: 1 of 152 No Filter Search



Text

Image editing

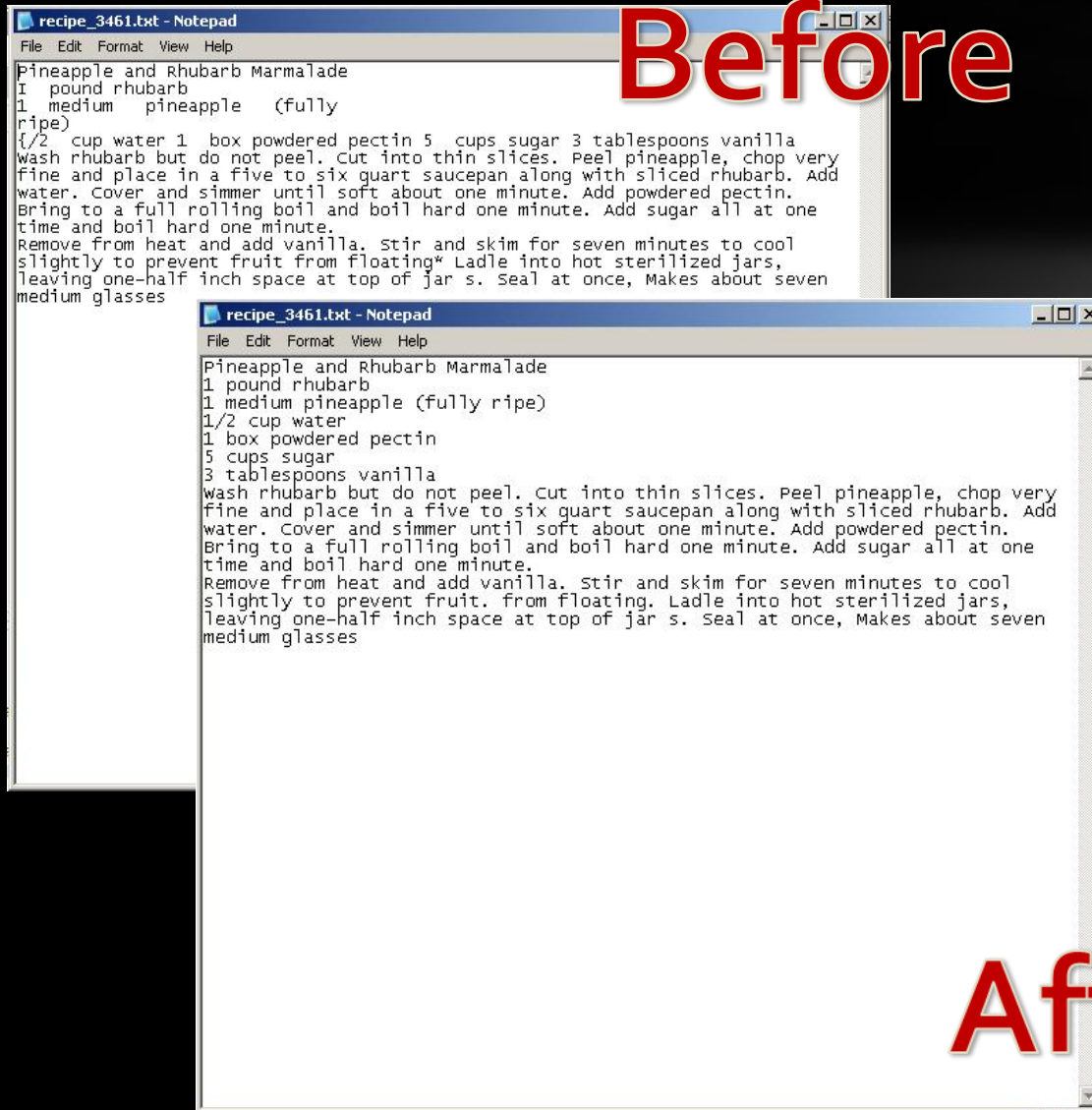
- Images need to be prepared before assigned to staff
- We use Photoshop but generally any basic image editor will do for minor edits

Before



After

Text creation



- Any OCR conversion needs to be done beforehand
- We use ABBYY Finereader
- CONTENTdm has an OCR component which works well for collections that need very little proofreading
- We save files as .txt to strip out unnecessary formatting
- Some files require much more proofreading than others
- The older it is, the more work it is likely to require

Database

The screenshot shows a web-based form titled 'Historic Recipes'. The form is organized into two columns of fields. The left column contains fields for ID, Title, Meal, Course, Ethnicity, Main Ingredient, Type of Dish, Seasonal, Recipe Source, and Transcript. The right column contains fields for Subject, Preparation Time, Newspaper, Date Appeared, Relation, Original Item ID, Type, Has Format, Collection, Language, and Submitting Institution. Most fields are text inputs, while Meal, Course, Ethnicity, Main Ingredient, Type of Dish, and Seasonal are dropdown menus. The bottom of the form features a record navigation bar with 'Record: 1 of 152', a 'No Filter' button, and a 'Search' button.

Field	Value
ID	1
Title	State Fair Cream Puffs
Meal	Snack
Course	Dessert
Ethnicity	American
Main Ingredient	Cream
Type of Dish	
Seasonal	Summer; State Fair
Recipe Source	Wisconsin State Fair
Transcript	
Subject	Cooking, American; Baking; Pastry
Preparation Time	
Newspaper	Milwaukee Journal
Date Appeared	1980-07-30
Relation	Recipe_0004.jpg
Original Item ID	Cream Puffs
Type	Text
Has Format	Recipe
Collection	Milwaukee Public Library Historic Recipe File
Language	English
Submitting Institution	Milwaukee Public Library

Record: 1 of 152 | No Filter | Search

- We use Microsoft Access for metadata entry
- There are advantages and disadvantages to using Access

Historic Recipes

ID: 1

Title: State Fair Cream Puffs

Meal: Snack

Course: Dessert

Ethnicity: American

Main Ingredient: Cream

Type of Dish:

Seasonal: Summer; State Fair

Recipe Source: Autumn, Christmas, Easter, Fourth of July, Halloween, Labor Day, Memorial Day, New Year's, Spring, State Fair, Summer, Thanksgiving, Winter

Subject: Cooking, American; Baking; Pastry

Preparation Time:

Newspaper: Milwaukee Journal

Date Appeared: 1980-07-30

Relation: Recipe_0004.jpg

Original Item ID: Cream Puffs

Type: Text

Has Format: Recipe

Collection: Milwaukee Public Library Historic Recipe File

Language: English

Submitting Institution: Milwaukee Public Library

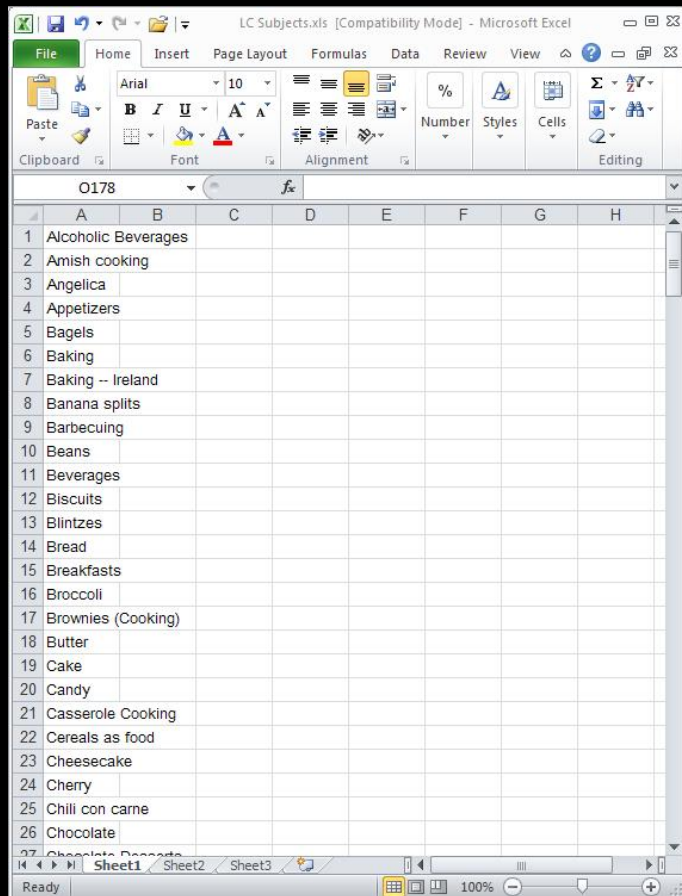
Record: 1 of 152

- Create drop down menus for controlled vocabulary lists
- Can't enter multiple values in a field from a dropdown list

Controlled Vocabularies

- Controlled vocabularies allow you to limit the terms that are used in certain fields.
 - Important for consistency when multiple people are working on the same project
- We use pre-existing controlled vocabularies i.e. LCSH, and create our own for local or specific subjects that don't already have one

Selected LCSH headings



LC Subjects.xls [Compatibility Mode] - Microsoft Excel

	A	B	C	D	E	F	G	H
1	Alcoholic Beverages							
2	Amish cooking							
3	Angelica							
4	Appetizers							
5	Bagels							
6	Baking							
7	Baking -- Ireland							
8	Banana splits							
9	Barbecuing							
10	Beans							
11	Beverages							
12	Biscuits							
13	Blintzes							
14	Bread							
15	Breakfasts							
16	Broccoli							
17	Brownies (Cooking)							
18	Butter							
19	Cake							
20	Candy							
21	Casserole Cooking							
22	Cereals as food							
23	Cheesecake							
24	Cherry							
25	Chili con carne							
26	Chocolate							
27	Chocolate Desserts							

Controlled vocabulary created in-house



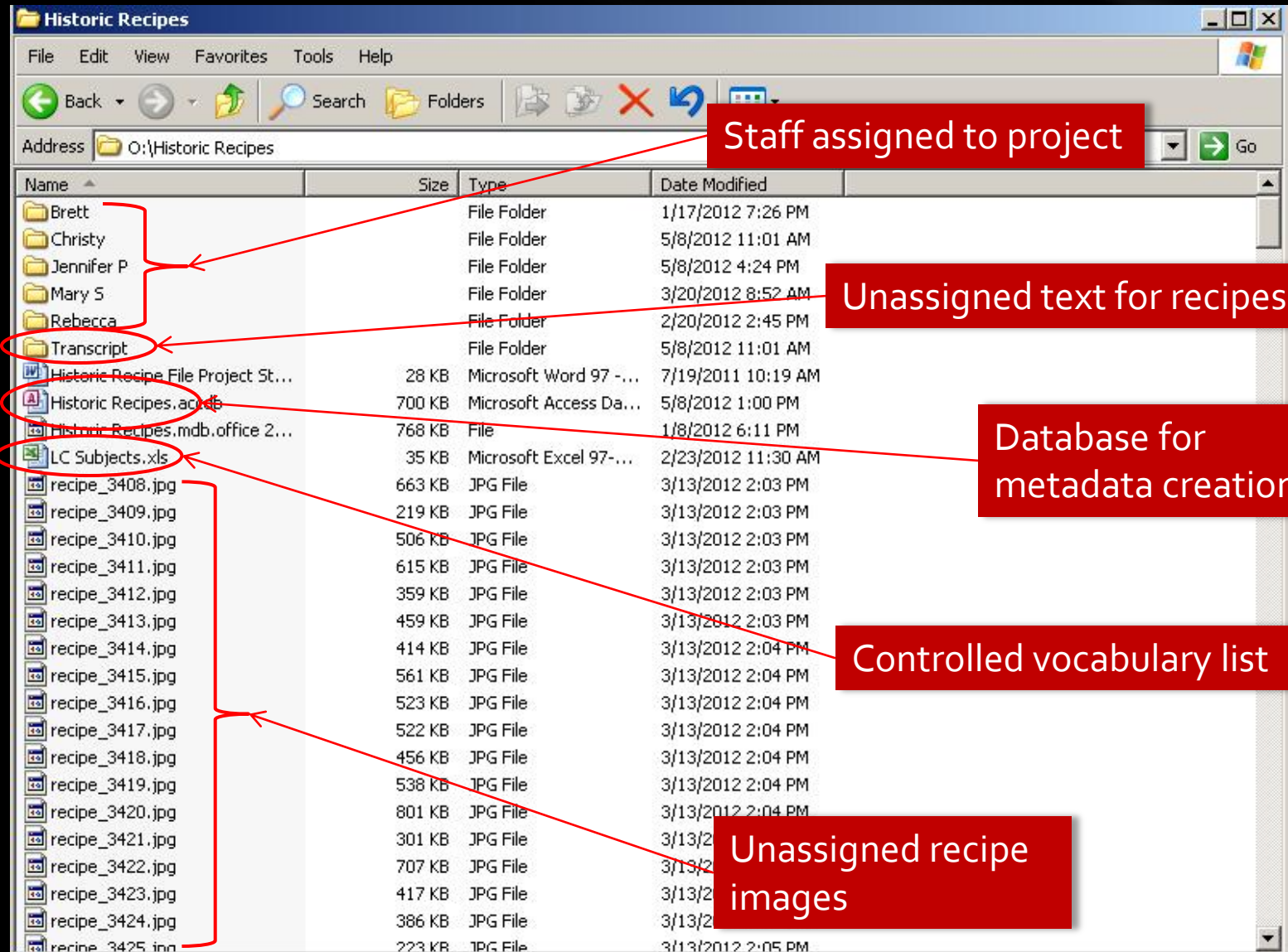
WWIBattlesCV.txt - Notepad

File Edit Format View Help

Adriatic Sea Theater
 Aegean Sea Theater
 Aigney-le-Duc Sector
 Aincreville USE Meuse-Argonne, Aincreville
 Aire River USE Meuse-Argonne, Aire River
 Aisne
 Aisne, Chemin des Dames
 Aisne-Marne
 Aisne-Marne, Bellevue Farm
 Aisne-Marne, Bois de Fere
 Aisne-Marne, Bois de Jomblets
 Aisne-Marne, Bois des Grimpettes
 Aisne-Marne, Buisson de Crenes
 Aisne-Marne, Champagne-Marne
 Aisne-Marne, Croix Rouge Farm
 Aisne-Marne, Fismes
 Aisne-Marne, Hill 230
 Aisne-Marne, Noroy
 Aisne-Marne, Ourcq River
 Aisne-Marne, Sergy
 Aisne-Marne, Seringes-et-Nesles
 Aisne-Marne, St. Thibault
 Aisne-Marne, Villars-sur-Fere
 Aisne-Marne Defensive USE Belleau wood
 Albert USE Somme, Albert
 Alsace USE Alsace Sector
 Alsace Lorraine USE Alsace Sector
 Alsace Sector
 Amiens
 Ancre River
 Ancy-le-Franc Sector
 Andevanne USE Meuse-Argonne, Bois d'Andevanne
 Andon Creek USE Meuse-Argonne, Andon Creek
 Angecourt USE Meuse-Argonne, Angecourt
 Aprement USE Meuse-Argonne, Aprement
 Archangel USE Russian Front - Archangel
 Ardennes Canal USE Meuse-Argonne, Ardennes Canal
 Ardeuil USE Champagne, Ardeuil
 Argonne USE Meuse-Argonne, Argonne Forest
 Argonne Forest USE Meuse-Argonne, Argonne Forest
 Argonne Forest Sector

Bringing it all together

- Once all the pieces are together create a place where staff can go to access them
- We organize the pieces into a main folder and then subfolders by staff name



Final steps

- Once the metadata is complete, the Access databases are harvested and converted to tab-delimited Excel files
- Images and metadata are parsed together while being imported into CONTENTdm
- Final quality control happens at this stage
- <http://content.mpl.org/cdm4/browse.php?CISOROOT=/histrecipe>

Some things to consider

- Create a project guide that takes staff through the project step by step
 - A good reference for them to consult if you're not available
- Consider creating a local style guide for your metadata fields
 - Increases consistency of entries in fields
 - Also good to consult if you're unavailable

What do I do if I don't have...

- There are other options beyond those gone over here, many which are much cheaper
- Determine delivery of images and accompanying materials to staff
 - Can be as simple as a flash drive with images loaded if no network is available
- Determine way of entering metadata
 - Google Docs has a spreadsheet with an option to create forms
 - <https://docs.google.com/spreadsheet/ccc?key=oArkFKY8gDS4wdDh6ZjZJQ3p1c2hleGgiYURINIR6VUE&pli=1#gid=0>