

Critical Reflection

The Schema for this assignment was created as a common document with my group mates Ana and Roxanne.

We started by noting down some generic elements that would form the structure of the schema. Initially our root element was 'corpus'. But we realized naming it 'document' made more sense since the CCP corpora consisted of documents of the Conventions, and because that's what we were going to input in our XML files. Most of these initial elements were intuitive. We also used the TEI crib sheet to help us get started with this list. Like title, body, page break, paragraph, name, and so on. Keeping the CCP principles in mind for using their materials, the 'type' attribute within our 'name' element has the option to specify if the name mentioned was an enslaved name, a chosen name, or an alias.

As we moved past the generic elements, our next step was to decide what documents to encode so that we could customize the other elements accordingly. After some research, we decided to pick the documents from a couple of exhibits which focused on the role of black women in the movement. Namely, 'BLACK WOMEN'S ECONOMIC POWER: VISUALIZING DOMESTIC SPACES IN THE 1830S' and 'WHAT DID THEY EAT? WHERE DID THEY STAY?' These two exhibits gave us a diversified range of elements. This led to the decision that the first two elements to be mentioned before any other would be 'exhibit' and 'CCP header'. These two elements help recognize where the documents were taken from the CCP website.

Since Mrs. Amie Long's menu consisted recipes within, a lot of the schema's elements are tailored around it. The tailoring became tricky when words like *enough*, *bits*, and *little* were being used as a unit of measurement. A part which caused a debate was words like *whole grains*. Does the *whole* here count as a quantity or a unit of measurement? In the end we decided that *whole* was better representation of the type of ingredient to be used rather than the quantity. An interesting aspect of creating a markup for the recipes was the *timer* element. This element marks not only the time mentioned in minutes, but also marks words like 'until done' and 'until brown'. The Boarding and Lodging exhibit was more focused on elements like name and locations. There were different types of locations mentioned in the document and it felt important to be able to identify those. So, our 'location' element includes the 'type' attribute which specifies whether the location mentioned is an address, state, or city.

It made sense to me to keep the definition of the structural elements consistent. So, I created a 'model_mixed_content' (MMC) which contained a compilation of elements that would be useful to mark up the text, rather than mark up the structure of the document. Elements like *emph*, *abbr*, *list*, *name*, *location*, etc., were included in MMC.

As I started working on encoding the documents, I identified some elements I had not thought of. Like a *lb* and *faded*. The schema remained a live document throughout the encoding process as I spotted other issues and fixed it accordingly. The nesting of elements was a delicate

process. When creating the structure before encoding, I realized I missed out on certain aspects like having the *caption* and *pagenum* inside the body, or having the *location* and *date* inside the *faded* element. Some of the other small details I missed were:

- Missing the `<text/>` in various elements which had an attribute.
- Decimals required in the 'quantity' element.
- A separate 'year' element.
- The page break element did not require numbers. But a separate page number element was required.

I dare to hope that the kind of encoding done would work well with an analysis tool. Identifying the popular units of measurements and quantities would help identify the context behind the recipes. Whether they are meant to suffice a family of 4 or a larger group, for instance. While a few elements are specific to the documents that I have encoded, I believe the schema can be used for other similar documents. It would certainly need more adjustment as and when other documents are being encoded, but it makes a good starting point.