

Jennifer Kaplan
Chris Klippenstein
Matthew Kumar
Vera Senina

Final Project
Indexing Ingredients: Using Ingredients to Navigate Making

Description of the Project

Our project tracks the relationships of recipe ingredients with recipe outcomes, as well as between multiple ingredients. (See our original proposal for definitions of "ingredient" and "recipe outcome.") We present these ingredient/ingredient and ingredient/outcome relationships in a variety of ways: through a newly programmed database that allows users to search for ingredients, through a [chart](#) that illustrates the co-occurrence of specific ingredients, and through ingredient profiles that highlight patterns of use. The database allows users to input their own search terms in either an ingredient- or result-capacity; a search for "egg" with the "ingredient" tag, for example, will return all the folio numbers in which "egg" appears as an ingredient. For this part of the project, we made use of the Making and Knowing Project's existing tagged terms; the work of identifying ingredients and results had already been performed. Our innovation was to contextualize these tagged terms as part of a searchable database. Matthew was our project's programmer and undertook all of the digital challenges.

In order to chart the co-occurrence of different ingredients, we first needed to simplify and normalize the results. The ingredients whose names included multiple tokens presented particular difficulty, as we needed to decide which ingredient names could be reduced to a single token and which should remain multi-token. Matthew generated a list of all the unique ingredients that appeared in the manuscript. Then Chris and Vera combed through the list of multitoken names (about [900 entries](#)) and tried to simplify them to their most basic form. For this step, we were joined by Jennifer, a later addition to our team who helped us normalize our list. We had to make many interpretive decisions about affinities between ingredients: for example, is "glair of egg" the same as "water of glair of egg"? If not, what might their important differences be?

In the end, we established a tentative system whereby we eliminated adjectives that were related to the processing of ingredients rather than their fundamental character. "Warm water" would thus become "water" because its temperature could be easily and reversibly altered. "Sandy earth," on the other hand, needed to remain "sandy earth" because the inclusion of sand made this ingredient fundamentally different from simply "earth." Vera created a [document](#) that provides a summary of difficult cases we have encountered and our treatment of those cases.

Of course, by simplifying some of these ingredient-terms for use in the chart, we naturally lost a considerable degree of nuance. We did not think this loss was fatal because the idea behind the chart was to generate a model — to use Dennis's favourite term — that could

illustrate the broad strokes of interaction between ingredients. We envisioned the chart as showing interactions between “water” and “eggs,” for example, without preoccupying ourselves in that moment with the question of whether they were crow eggs or chicken eggs. Vera created the chart using data provided by Matthew.

Finally, our prototype [ingredient profiles](#) re-introduce the nuance that our chart omits. Each profile offers detailed information about the type of ingredient, its various manifestations throughout the folio (re-expanding the chart category "earth" into the subcategories of "burned earth," "fat earth," and so on), the folios in which it appears as alternately ingredient and recipe, and any interpretive questions or problems that arose as a result of looking at the ingredient subcategories side by side. In terms of the digital edition, we imagine an interface in which users would be able to click on the different ingredients in our chart to learn more about them. Chris wrote the ingredient profiles on "earth" and "egg"; Jennifer wrote "oil" and "smoke"; Vera wrote “varnish”.

Insights About Course Content and Ms. Fr. 640

Although we decided to organize the folio according to ingredients, it quickly became clear that this method would not be equally suitable for all the entries in the manuscript. There are many entries that are not necessarily traditional recipes (which produce an object), but how-to guides (which produce an effect), and these how-to guides may or may not use ‘ingredients’ as we defined them. We sometimes noticed a subtle distinction between the titles of object-producing recipes and the titles of how-to guides, and we wonder whether it would perhaps be possible (or useful) in the future to separate descriptions of object-producing recipes from descriptions of effect-producing recipes.

Extracting the recipe outcomes automatically also posed a problem, which we could only resolve by relying on the approximation that the recipe title equals the projected outcome, since not all recipe titles contained names of materials (many how-to guides did not). However, this approximation later appeared to be problematic in several ways. For example, the assumption that a material present in the title equals the outcome of that recipe could often be misleading, since the recipe outcome in those cases could also be the elimination of the mentioned ingredient and not instructions on how to make it. Another issue is that the titles in the manuscript do not always present an accurate representation of the projected outcome; interpretation and careful reading are often still necessary. At this point it appears that classification of titles may require human annotation (or, at least, human verification) -- and this seems to be the case with all tasks that require higher precision. One of the things we learned, then, is that the technical side of our project was still heavily dependent on human interpretation, input, and formalization. The search and co-occurrence functions in our project, for example, are limited by which ingredients are human-tagged. To deal with the variance in ingredients and interpretation, an entire other document had to be made by Chris, Vera, and Jennifer to identify terms to find and replace with others.

All of these findings support the structure of our class together, which relied on a mixture of the digital and the artisanal, a combination of work with computers and work done by hand. Our project only exists because we were able to bring programming skills together with our capacity for careful, painstaking, analogue work -- in doing so, perhaps we inadvertently found an echo of the writer-practitioner's own use of tool-supplemented and analogue work.

From Proposal to Execution

We remained faithful to the central tenets of our proposal, though we did veer away from the "ingredient pipeline" that we had initially found attractive. Our vision of the project was to make a search tool that also incorporated content analysis, and while we feel that we have reached this goal, much of our tool is still a prototype rather than a completed product.

In a perfect world, for instance, our chart would include information about every ingredient, rather than merely a selection. (We made that selection through the unscientific process of choosing ingredients that we thought would be reasonably common, or ingredients that we had found particularly interesting.) We would have ingredient profiles for every ingredient, and we would expand these profiles to include (1) provenance information about that ingredient (providing potential for crossover with the mapping group), (2) images of that ingredient (offering a point of connection with our fellow NLP group and their chromatic index), (3) a list of ingredients with which a specific ingredient commonly appears, and (4) notes on the process of transformation between ingredients, for example, which transformations are enacted through heat.