

## Crowdsourced Digitization

Due: Friday, October 22, 5pm. Points, per the syllabus: 1.

Last week we looked at digital datasets. For this week, we will actually contribute to these.

There are many crowdsourced history projects out there. The purpose of this is to get you some more experience looking *closely* at a historical dataset.

1. Choose a crowdsourced historical data project.

I recommend choosing one from the [Zooniverse history projects](#), but you can pick your own as long as it targets a paper database of the sort we have been discussing. Libraries like the [Library of Congress](#) maintain their own lists.

## Crowdsourcing digitization

Try **not** to pick something is primarily about transcribing longer stretches of text. As a rule of thumb, if there are *sentences* it might be too long (as opposed to labels). If the paper is divided into boxes, though, you're probably still OK.

2. Spend 30-60 minutes contributing to that project. (Build in time before this to create an account, read training materials, etc.).
3. Write a short (500 words) response in the assignments tab of the course about the experience.
4. What did you transcribe? What kind of data? Who seems to have sponsored it?
5. What is the value of the data that you transcribed? Is it the same as the sponsor would want?
6. How difficult was it (technically, emotionally, whatever) to do the transcription?
7. What sort of things that existed on the paper did you have to omit?
8. What sort of the things were not on the paper that you would have expected to be?
9. What does it indicate about what kind of data might be "good for digitization" that you wouldn't have necessarily known before?