# Researcher Bi-Weekly Reporting Template

The purpose of this document is to help prompt reflection and capture the progress, barriers, and needs you have as you undertake your project. We’ve included some preferred and suggested prompts in each section below. This document is to be completed every two weeks and submitted according to the reporting schedule.

Please append the filename with the date of submission and your initials (ex: CCHC-Biweekly-Reporting-Template-**2021-06-01**-**od**.docx). Once completed, please submit to [**LC-Labs@loc.gov**](mailto:LC-Labs@loc.gov)and CC Jaime Mears ([jame@loc.gov](mailto:jame@loc.gov)).

**Name: Lauren Tilton**

**Reporting Period: 2: 6/14/2021-6/25/2021**

### No Work Completed

We anticipate that there may be periods in your project during which you make little progress. In this situation, we ask you to check the “Nothing to report” box. After the “Reason” prompt below, please briefly indicate why there was nothing to report. Where possible, we’d still like to learn more about the reasons behind any unexpected events or barriers you encounter. You can share additional details in the “Obstacles/Needs” section before submitting this report.

Nothing to report

Reason:

### Activity and Progress

*We’d like to hear about your progress over the last couple of weeks. Please feel free to expand on the questions below.*

1. **What did you and your team work on over the last two weeks? Why?**

*We’d love to hear about your decision-making, the technologies (ie. cloud services, hardware, and software) you’ve been using, why you’ve adopted them, modifications or hacks you’ve had to create, prototyping and testing, and anything you’d like to share about your current in-progress deliverables.*

We worked on connecting the images together with the metadata. This involved downloading the images according to the metadata, specifically figuring out what the URLs are. We decided to download the highest resolution jpeg file.

1. **What are your plans for the next two weeks? What excites you about those plans?**

Next week, we will begin applying image embedding algorithms to the collection.

### Obstacles/Needs

1. **What barriers, if any, have you encountered over the past two weeks? Were these barriers expected?**

*Examples: unexpected costs; API limitations; organizational barriers; data issues; technical gaps or challenges; collections knowledge.*

The challenge was trying to figure out which image file to download. There are different file types as well as sizes for each file type. There isn’t documentation, that we could find at least, that explained why there are different file types and the decisions behind the scans. We kept asking ourselves: Which resolution to use? Are the TIFFs better than JPEGS?

1. **What would resolve these barriers (or would have resolved them, in retrospect)?**

*We’d love to know what you need help with! This is an opportunity for LC staff to learn from you about how to best support this work going forward.*

We looked at some examples and decided to use the largest JPEG. The TIFFs are huge and this means challenges for storage and analyzing. One thing we did was take a few of the TIFFS and convert them into JPEGS. Once converted to JPEG, it was a similar size to the other JPEG versions. Therefore, it appeared that the largest JPEG was either a converted TIFF or at least close to it. The compressed large JPEG seemed sufficient for our purposes.

High quality JPEGs works well for our purposes. Compressed images files are fine for CV as long as the file is large enough. Our rule of thumb is something close to HD is good enough for computer vision. It does depend on what you are looking at though. If you have a photo with a lot of people and want to identify a specific person, the file might need to be bigger. So,

It would be nice for there to be documentation for each collection explaining the scanning decisions. For example, is the JPEG just a converted TIFF? Documented decisions like this would be helpful.

1. **What changes, if any, do you anticipate to your proposed timeline? Why?**

*Do you expect any significant blockers to your progress?*

None at this moment.

### Reflection

1. **What happened over the last two weeks that was particularly thought-provoking?**

*Examples: surprises; places where help from Library staff was critical; incorrect assumptions; things you learned; things that, in retrospect, you wish you’d known.*

It was interesting to think about histories and processes for digitization.

1. **Anything else you’d like to add?**

Olivia, Alice and I discussed the digitial file issue during our 7-9 meeting. I appreciate their offer to look into the digitization information about the collections, and to see if we should meet with specific folks who are familiar with the decision making behind the file types and sizes. I shared that I often find that the archivists and other colleagues who were involved in processing the collection have all kinds of knowledge about the collections that isn’t documented anywhere. It’s always great to learn from them about the collection! It would also be great if there was a section that outlined all the digitization decisions.