# 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:**

**Reporting Period: 12: 11/1/2021-11/12/2021**

### 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.*

A major part of this week was meeting with PPOC. It was a great opportunity to get the perspective of the team. We presented on the project. One aspect that I worked on was developing slides the went more in-depth about each of the CV techniques being used. Based on feedback from a previous meeting, we also provided an example of how CV helped with more formal readings of the photographs. This extended the payoff of these methods. They aren’t just about access and discovery across the collection but could also help us “view” an individual image or small set of images in a new way. The example I drew on was the work of Esther Bubley and her work in DC. Using image segmentation, we can get a sense of her ways of seeing. My use of the term “ways of seeing” is intentional here because it draws on John Berger’s book by the same name. It’s a foundational work in the study of visual culture because he demonstrates how we develop and circulate certain ways that we create art, in other words a certain kind of gaze, that in turn reflect cultural and social values. The relationship between sky and ground brought my attention to the way that she looks up to buildings in DC, literally looking up and conveying their monumentality. There is a reverence for DC that is convey through how she gazes at the built environment of the capital.

The conversation that ensued was very interesting. Discussions included:

* Discussing the relationship between digitization and computer vision. For example, I discussed how borders and targets become an issue if a part of the scan. This led to some questions and ideas about digitization. I shared that what is best for it to be consistent and documented. So, if a whole collection is an over-scan or has targets, then we can write some code to crop the image accordingly (assuming all the sizes are the same, so all 4 x 6 prints, safety film, or another form).
* They are still working through digitizing parts of the collections that I’m working with. Given contracts and what’s being digitized by size or type, it is not uncommon for a part of a collection to be digitized. I shared that it would be nice if this is documented. For example, they have done all the nitrate for FSA and now working through the sheet film. It would be great if this was documented somewhere.
* Suggestion that the technology, i.e. the kind of camera, is really important to understanding these collections. There was interest in the visualizer somehow conveying this kind of information. I’m still thinking through this. It might be a matter of just making sure there are links to the relevant LOC pages that give more background and context.
* There was some concern about AI. One member of PPOC had worked with some CV methods and was horrified by the results. She had concerns about the approach, which we discussed. It was nice to have on the table this issue, which is very much a pressing one. We discussed how one approach is to map concepts onto different concepts as a strategy.
* There seemed to be general interest in ways to facilitate browsing, which people liked in the current version of the visualizer. We also asked about the metadata question and there seemed to be consensus that the metadata be included.

With a little distance from this meeting, a major take away was two parts.

1. One is that there is a lot of interest in context. This is a challenge for the visualizer because it is designed more as an exploratory interface that suggests possible connections rather than argumentative or very interpretative interface, at least to the degree that I was hearing emphasized. The interface isn’t trying to say that all connections are meaningful or worth studying. Rather, it’s suggesting/ recommending *possible* connections that will need to be further explored, contextualized, and studied. I’m thinking we will need to make this very clear. We are thinking we will add an About page that outlines what ADDI does and *does not* aim to do.
2. The second was about digitization and process. The rise of researchers using CV means that the relationship between the digital image, the digital file, and the researcher is changing. Before, the person looked online at the image and mentally threw out the pieces of the images that were evidence of digitization such as targets. Most CV methods look at the whole image to find the feature that the algorithm has designed to look for. Neural nets learn from the entire image. The targets, border of the negative, paper mounts, frame, etc. all become “the image” to the computer. CV requires certain file quality/size as well. Old scans with low DPI are often illegible to CV algorithms. The way that CV “sees” images means that we may have to account for another way looking at digital files. To facilitate computational methods of images, this may mean adjusting or adding to the digitization process for image collections.

The PPOC meeting also led to some other ideas that we discussed during the 1:1 on Nov 9. They included:

* The idea for a page online that tracks digitization. It could include what was just digitized, what’s being digitized, and what’s next. It should also include the project teams. I thought it would be cool because it would be a great resource for researchers as well as a nice way to show the incredible amount of work that is underway at LOC. The rate and amount of digitization is incredible! It is something that I didn’t understand til this experience. The public should know this!
* Another idea that emerged was a check list of questions to ask/ look into before working with the data. I know I’d love this for teaching with data as well! For example, we talked about how Bain is all that LOC has, but it is known that the bulk of the work burned up in a fire. So, the collection is all that LOC has but it is not comprehensive. This changes what questions one can ask of the data. This isn’t the best way to frame these but want to offer some ideas! A list could include items such as:
  + What does “whole” mean for this collection?
  + To what degree is this the whole collection?
  + Explore the metadata. Is it clear what each field means? If not, consider emailing the curator.

The other meeting during the time period that led to interesting questions/ideas was the Open Office Hours on 11/9. We discussed UX. One aspect that came out of the discussion that we are working through is the design of the visualizer. The issue came up again of authority and the visual rhetoric of digital projects in general (not just visualizations/graphs). When a site is sleek and following certain visual conventions, it conveys authority (at least in US visual culture). For example, Helvetica is more formal while Comic Sans is considered childish. High saturation colors are considered more playful (i.e. kids cartoon site like Nick.com) or consumer oriented (i.e. McDonalds) while gray scale is associated with documentary/realism/seriousness (i.e. NYTimes, The Atlantic, LOC Search). One idea was to keep the visualizer looking like a prototype as a design choice. One thought is that by conveying visually that it’s a prototype, it will signal caution to the user, a productive distrust of the results. I’m trying to think through what that looks like. I’ve been percolating on this.

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

With the end of the semester coming and then moving the weekend of 11/19, I don’t anticipate having a lot of time to work on this. It’s usually the week before Thanksgiving Break that is the busiest and things begin to lighten up after Thanksgiving.

### 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.*

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.*

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

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

We are going to need a few weeks in December to work on the final visualizer.

### 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.*

I focused a bit on this in Question 1. I’m still percolating on the design elements of the visualizer and thinking through this idea of keeping the prototype looking prototypey (I think this is a new word ☺ ). I’ve also been thinking a lot about digitization workflows after the conversation with PPOC, and how researchers who want to use images as data might need the data in a different way than is standard/common practice in image digitization. This seems like it would also be an interesting article/ piece for a journal (if others haven’t written about it already!). I’d be interested to work on such a piece if there was interest within LCLabs and/or PPOC.

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

As always, thank you to everyone for all their work!