CALIFORNIA PHENOLOGY TCN – QUARTERLY REPORT – FEBRUARY 2021

Assembled by Katie Pearson, 1 February 2021

PROGRESS IN DIGITIZATION EFFORTS:

Figure 1 shows our progress in imaging, transcribing, georeferencing, and phenologically scoring the target specimens for the original 22 CAP institutions, explained more in detail in the following sections.

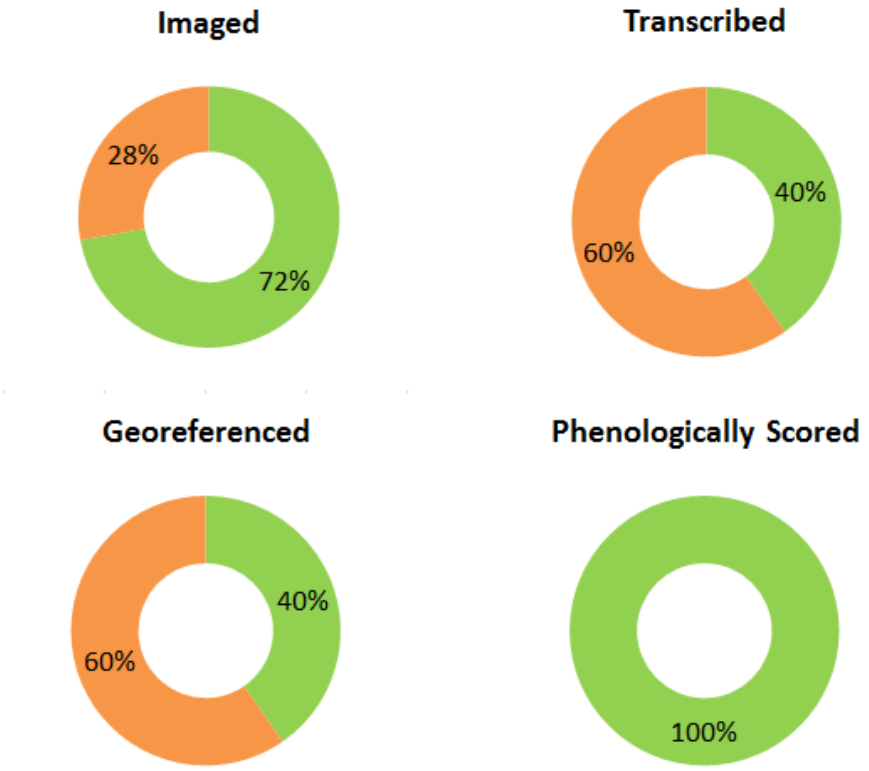

Figure 1. Progress in each of our four major digitization goals. Note that totals represent the original goals of the CAP grant: 902,400 specimens imaged and phenologically scored, 300,000 specimens transcribed, and 300,000 specimens georeferenced. Additional specimens to be digitized by the PEN grant are tabulated in the PEN section below.

TRANSCRIPTION

An estimated 120,000 specimen records have been transcribed across the CAP Network since the beginning of the project. This is approximately 40% of the goal number of transcriptions.

Transcription has largely been accomplished by student volunteers working for credit (see Education and Outreach section) and through Notes from Nature expeditions. One NfN expedition was completed in November 2020, and we have launched two new expeditions in addition to the two existing expeditions.

GEOREFERENCING

We have georeferenced approximately 121,048 specimen records, which is 40% of our goal. Georeferencing is being conducted by trained staff and students at HSC, OBI, and SD, and by naturalist volunteers are part of the "100 Club." The CAP 100 Club currently has 26 active members who have collectively georeferenced over 6500 specimens since September 2020.

After switching to live management of their collection in the CCH2 portal, RSA was able to merge a file of previously-created georeferences into their specimen records, which added over 139,000 georeferences to the CCH2 portal.

IMAGING

Eleven of our 22 herbaria (50%) have completed their imaging goals (Figure 2). Of the remaining herbaria, four have been able to resume imaging since the COVID-19 shutdowns. The other herbaria have used this time offsite to process images, georeference specimens, and transcribe specimens. Figure 2 shows the current state of CAP imaging as of February 1, 2021.

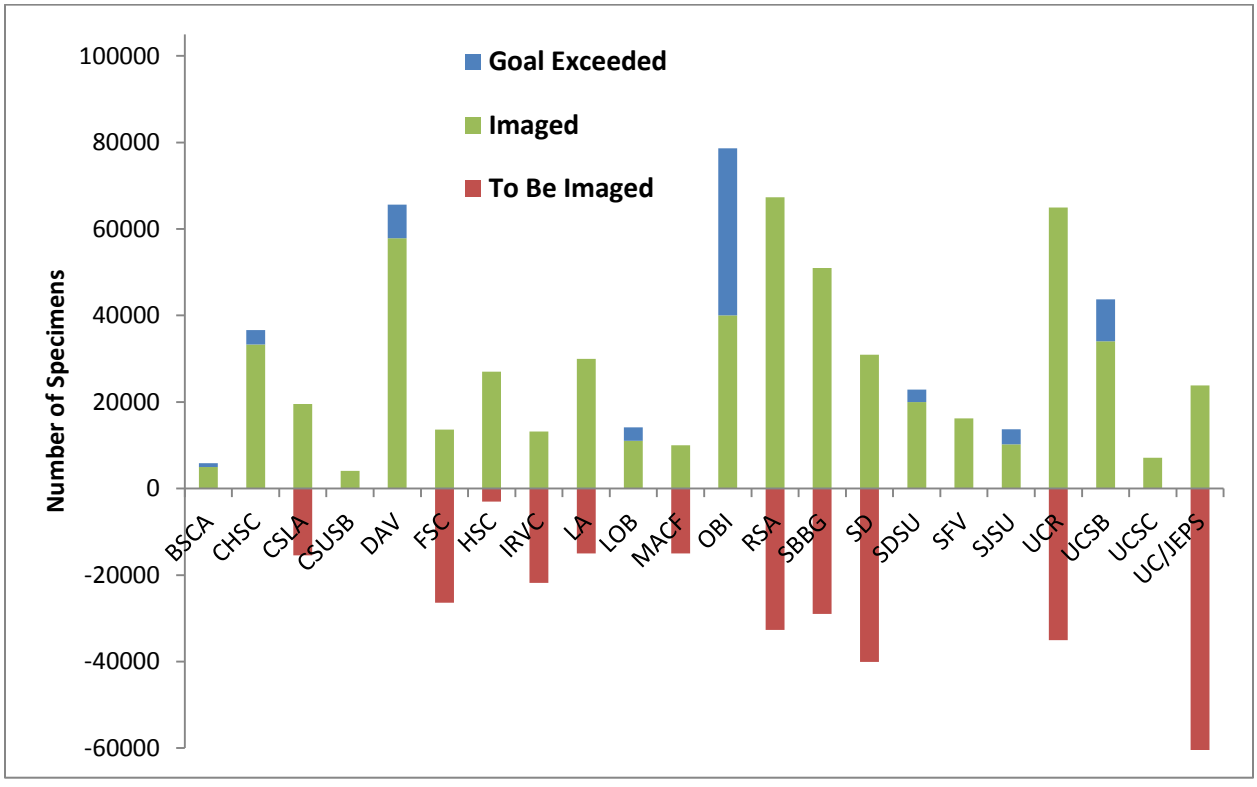

Figure 2. Herbarium specimen imaging progress. Green portions represent the number of specimens that have been imaged, while blue portions represent the number of specimens that have been imaged beyond the expected target specimens. Red bars below the zero line indicate the number of target specimens that have not yet been imaged.

PEN PROGRESS

Of the seven CAP institutions funded through the 2020 PEN grant, three (OSC, SD, UNLV) are currently imaging specimens. One (CDA) was delayed as we awaited approval to convert a subcontract to a sub-award. One (SFSU) has been delayed by COVID-19-related limitations. The remaining two will begin imaging later in 2021 (SHTC) and 2022 (PUA), as previously scheduled. Figure 3 shows the current imaging progress at PEN institutions.

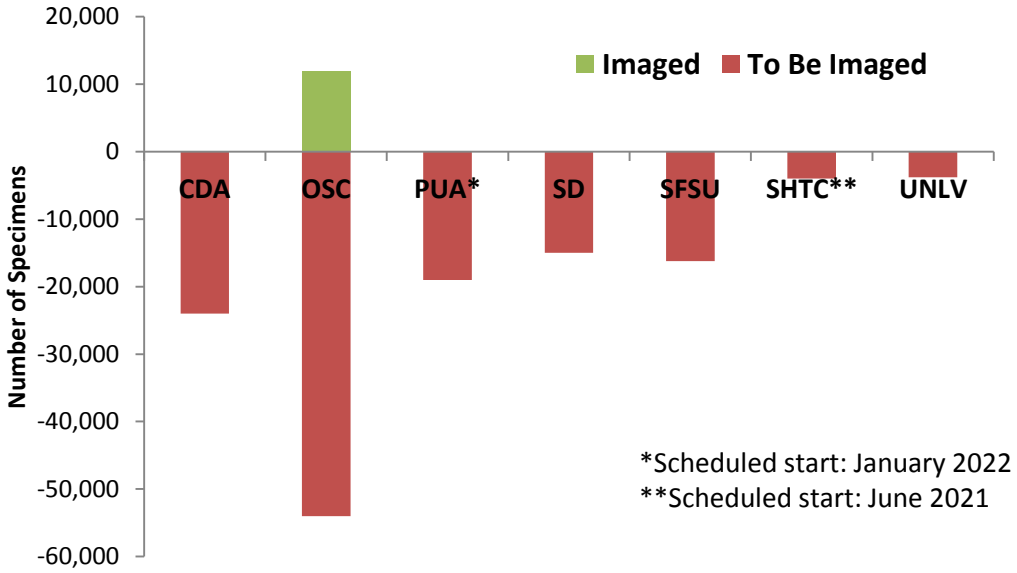

Figure 3. Herbarium specimen imaging progress for the seven PEN institutions. Green portions represent the number of specimens that have been imaged, while red bars below the zero line indicate the number of target specimens that have not yet been imaged.

The PM has continued to keep in close contact with the PEN institutions and ensure progress toward imaging goals. The PEN partners are included in all CAP-wide communications and meetings.

The data from all PEN institutions are now in CCH2 with the exception of PUA. The PM and PI at PUA are working to convert a previous RBase database into a file that can be imported into CCH2.

SHARE AND IDENTIFY BEST PRACTICES AND STANDARDS (INCLUDING LESSONS LEARNED)

Last quarter we reported the development of a tool that extracts georeferences for duplicate specimen records (https://doi.org/10.5281/zenodo.4300889). After adding the Oregon State University and Southern Oregon University collections to CCH2, we ran this script on all of the CCH2 specimens that were collected in Oregon. As a result, we identified over 2800 records that could be georeferenced from duplicate data. These georeferences were sent to the managers of their corresponding herbaria and uploaded into CCH2 when approved.

As described in the previous quarterly report, another strategy we employed to advance georeferencing is launching the "100 Club". Members of the 100 Club are naturalists, botanists, and other professionals who volunteer time to georeference specimens in CCH2. We currently have 26 active members of the 100 Club, and they have georeferenced over 6500 specimens since September 2020. On February 1st, we held a group meeting to build community and encourage the exchange of strategies for more efficient georeferencing. We are continuing to recruit members from the many naturalist networks across the state.

IDENTIFY GAPS IN DIGITIZATION AREAS AND TECHNOLOGY

We are continuing to develop the functionality of our CCH2 data portal. Currently, we lack a way to search by, download, and visualize specimens' phenological scorings, but we are working with Chris Tyrrell at the Milwaukee Public Museum to develop these tools, which are nearly complete.

SHARE AND IDENTIFY OPPORTUNITIES TO ENHANCE TRAINING EFFORTS

We have continued to recruit and train new members of the 100 Club to georeference in CCH2. We currently have 26 active members of the 100 Club, and they have georeferenced over 6500 specimens since September 2020.

During the fall quarter, Cal Poly led an online herbarium digitization course in which 65 students at 6 institutions learned how to georeference and transcribe specimens. Drawing on the lessons learned and the feedback received from students last quarter, we are leading a similar, yet smaller course for the winter 2021 quarter. This course meets synchronously, once per week for 2 hours and is composed of 32 students from 9 institutions. Five students are returning students from the fall 2020 course, which allows us to split into smaller breakout rooms during digitization tasks to facilitate more interaction and asking of questions. The 32 students are now trained in specimen transcription in Notes from Nature, and we will soon transition to data entry directly in CCH2. As part of the course, the students also give a short presentation on a scientific paper that reports using herbarium specimens in research.

PI-Liston at Oregon State University is leading a course of 12 OSU students in georeferencing Oregon specimens from CCH2. The PM has helped to facilitate the use of the GEOLocate Collaborating Georeferencing (CoGe) platform to this end, and the class is acting as somewhat of a beta test for future collaborative georeferencing efforts for the CAP Network. As part of this process, the PM also developed a user protocol for using CoGe for georeferencing, now posted on our website: https://www.capturingcaliforniasflowers.org/uploads/1/6/3/7/16372936/georeferencingincoge.docx

SHARE AND IDENTIFY COLLABORATIONS WITH OTHER TCNS, INSTITUTIONS, AND ORGANIZATIONS

PM Pearson has been hired by Arizona State University as the portal manager for the GLOBAL TCN bryophyte and lichen portals. This will continue to increase Pearson's Symbiota skills and provide an easy pathway for any California institutions that have not yet added their bryophyte or lichen data to

their respective portals. Pearson will continue to act as PM in the full capacity of the position for the remainder of the CAP TCN grant (through late July 2022). Pearson has also been able to advise the GLOBAL project manager and georeferencing manager on reporting, communication, and imaging- and georeferencing-related workflows.

Lead PI Yost and the PM presented at a UC Naturalists instructors' meeting on December 2nd, 2020. This presentation described the tools that are available to get UC Naturalists involved in digital specimen data, and we disseminated a document that outlines the potential options. We are continuing to collaborate with the UC Naturalists team to engage more of their membership in digital specimen data.

Specimen data and images from the Oregon State University (OSC) and Southern Oregon University (SOC) herbaria were added to CCH2 as snapshots. OSC manages their data in a local instance of Symbiota (the Oregon Flora), while SOC manages their data in the Consortium of Pacific Northwest Herbaria portal.

With organization from CAP leadership, the Consortium of California Herbaria approved a new MOU with Calflora to facilitate ingestion of CCH2 data into Calflora. The CCH and Calflora have not historically had a highly collaborative relationship; thus, these efforts have been pivotal for both organizations. During this process, Calflora shared a list of records that had previously been excluded from CCH → Calflora imports due to potential errors, and we have used this list to correct these errors, where possible.

CAP leadership is in communication with members of the Endless Forms TCN in an effort to build a community understanding of data protection for sensitive species.

The PM also met with Tiana Rehman and Diego Barroso of the TORCH TCN to discuss using the GEOLocate Collaborative Georeference (CoGe) platform for their georeferencing goals. The PM shared the newly-developed CoGe protocol with Rehman and Barroso, who will provide feedback for the further improvement of the protocol.

The PM reached out to leadership at the California Geography Alliance and the California Geographical Society about establishing a partnership to engage their membership with specimen georeferencing.

SHARE AND IDENTIFY OPPORTUNITIES AND STRATEGIES FOR SUSTAINABILITY:

The PM is being trained to query and make edits to the CCH2 database through the MySQL Workbench to enable more functionality than is currently available through the front-end interface.

We are encouraging PIs to lead herbarium-focused independent research or internship courses at their own institutions, and the PI is supporting their efforts in doing so with planning and training. These 1-3 credit courses will be a sustainable way to keep students working in the herbaria and contributing to digitization and collection curation.

SHARE AND IDENTIFY EDUCATION AND OUTREACH (E&O) ACTIVITIES:

The PM shares updates on the project and phenology-related news via the network Twitter account (@CalPhenologyTCN).

We are now aggregating CAP-related educational resources on a new page of our website: https://www.capturingcaliforniasflowers.org/educational-materials.html

One blog post was written and published to the CAP website: https://www.capturingcaliforniasflowers.org/blog-recap.

Similarly to last quarter, Cal Poly is leading an undergraduate internship course involving students at multiple CAP institutions. The course began in early January and currently involves 32 students from 9 institutions (Cal Poly, CSU San Bernardino, San Francisco State, Pacific Union College, UC Irvine, UCLA, UC Riverside, UC Santa Barbara, and UC Santa Cruz). There the course meets synchronously on Tuesdays for two hours and works on transcription in Notes from Nature. Students are also expected to give a presentation on a research paper that uses herbarium specimens. The progress of the course, as well as a gallery of interesting specimens, is maintained on the course website: https://www.capturingcaliforniasflowers.org/herbarium-digitization-course.html

PI Yost is also leading a research course at Cal Poly similar to our spring 2020 course, assisted by the new CAP-funded postdoctoral student (Natalie Love), a graduate student at UCSB (Tadeo Ramirez Parada) and the PM. In this course, students conduct original research using newly digitized data. They score the phenological status of specimens of a focal taxon, use models in R to analyze the effects of climate on the timing of phenological events, and write and research paper and create a research poster to summarize and disseminate their results. PI Mazer (UCSB) is sitting in to the course and plans to run the course at UCSB during the spring 2021 quarter.

PI Waselkov (CSU Fresno) and PI Liston (Oregon State University) are separately leading independent research courses that involve digitization. Waselkov's group works on transcription and meets weekly to discuss their progress and research regarding herbarium specimens. Liston's group is co-led by the PM and is focused on georeferencing specimens from California herbaria that are located in Oregon. This course meets synchronously (optionally) and students work synchronously together in small groups on georeferencing specimens in assigned Oregon counties.

One Notes from Nature expedition was completed in November 2020, resulting in 1030 specimens from UC Irvine being fully transcribed. The data will be brought into CCH2 upon the return of the UC Irvine collections manager from temporary leave (March 2021). We launched expeditions for two new institutions—UCLA and Humboldt State University—in November 2020 and January 2021, respectively.

As part of the BIOME institute, the PM created a short video on the topic of "How and When to Cite Biodiversity Specimen Data," now posted to the CAP TCN YouTube channel: https://www.youtube.com/watch?v=z7E6VvtNvpY.