

## CALIFORNIA PHENOLOGY TCN – QUARTERLY REPORT – FEBRUARY 2019

*Assembled by Katie Pearson and Jenn Yost, February 5, 2019*

### **Progress in Digitization Efforts:**

All institutions have been trained to image (for those not sending specimens to other institutions for imaging) and use the CCH2 portal as of February 1, 2019. Cal Poly (OBI; lead institution) hired 11 student imaging technicians and one supervisory student. Other institutions are also recruiting and/or hiring students, barcoding specimens, and beginning to image.

Katie Pearson (project manager; PM), Jason Alexander (data manager; DM), Jenn Yost (lead PI) and Ed Gilbert (Symbiota developer) have weekly video conferences related to portal development and data migration.

All live collections in the grant have been migrated into CCH2, our Symbiota-based data portal, reflecting 1,772,225 specimen records, 56% of which are georeferenced and 3% of which currently have images in the portal.

In the future, digitization progress will be reported in terms of new images created. Because this is the first report in which all institutions are uploaded in the portal and many institutions are still in beginning stages of imaging, we provide raw numbers in Table 1 and comment on the statuses of certain institutions.

**Table 1** Number of total specimens, number of georeferenced records, and number of imaged specimens for each institution in the CCH2 portal as of February 5, 2019. Shaded rows indicate institutions that have been actively imaging and successfully uploading images to the portal. Non-shaded rows indicate institutions that are currently preparing for imaging or have not been able to upload images. A single asterisk (\*) indicates institutions that have been imaging but have not been or have been unable to upload images to the portal, and therefore the numbers of images for these institutions do not reflect the true number of images created. A double asterisk (\*\*) indicates the institution that has not been uploaded into the CCH2 portal to date but did provide the number of images created.

<b>Collection</b>	<b>Specimens</b>	<b>Georeferenced</b>	<b>Imaged</b>
<b>BSCA - Colorado Desert District, California Dept. of Parks and Recreation</b> (sending to RSA)	3618	2848	0
<b>CHSC - Chico State Herbarium</b>	116798	95341	0
<b>CSULA - Cal State LA Herbarium</b>	0	0	0
<b>CSUSB - CSU San Bernardino</b> (sending to RSA)	3981	2585	0
<b>DAV - UC Davis Herbarium</b>	108571	80804	1*
<b>FSC - Fresno State Herbarium</b>	3379	5	2018
<b>HSV - Humboldt State University Vascular Plant Herbarium</b>	73108	12765	0

<b>IRVC - University of California, Irvine Herbarium</b>	15215	4118	*
<b>LA - UCLA Herbarium</b>	24511	2959	0
<b>LOB - CSU Long Beach Herbarium</b>	6883	2713	7*
<b>MACF - MacFadden Herbarium, Department of Biological Science, CSU Fullerton</b>	595	6	0
<b>OBI - Robert F. Hoover Herbarium, Cal Poly State University</b>	55073	30178	1050
<b>RSA - Rancho Santa Ana Botanic Garden Herbarium</b>	612870	143582	*
<b>SBBG - Clifton Smith Herbarium, Santa Barbara Botanic Garden</b>	129621	105776	52
<b>SD – Herbarium, San Diego Natural History Museum</b>	239037	191021	*
<b>SDSU - San Diego State University Herbarium</b>	21753	16685	545
<b>SFV - CSU Northridge Herbarium</b>	15122	11217	8838
<b>SJSU - Carl W. Sharsmith Herbarium, San Jose State University</b>	15761	5353	*
<b>UCJEPS - UC/Jepson Herbaria, UC Berkeley</b>	2**	0**	17432
<b>UCR - UC Riverside Plant Herbarium</b>	269826	252531	0
<b>UCSB - UC Santa Barbara Herbarium</b>	44568	20561	41929
<b>UCSC - UC Santa Cruz Herbarium, Kenneth S. Norris Center for Natural History</b>	11580	6742	235

**Share and Identify Best Practices and Standards (including Lessons Learned):**

All protocols are being edited and re-issued as new efficiencies are discovered throughout the digitization process. The PM is planning a webinar for March or April in which institutions “swap stories” about tips and tricks they have found to make the process run more smoothly.

The PM continues to monitor created images, solicit feedback from member institutions, and adjust recommendations as necessary to create the best possible images. For example, in response to concerns about dull coloration of images, we now recommend increasing exposure compensation at time of imaging. Similarly, Susan Mazer (UCSB PI and phenological researcher) reported that images of about 5-10 MB were best for viewing reproductive structures of specimens, and the recommended size of processed jpegs was increased to meet this standard.

We will initiate discussions with our review board this coming quarter to establish the phenological standards and terminology with which we will codify phenological data. Ed Gilbert is continuing to work on ways to collect and store these data in the CCH2 portal.

**Identify Gaps in Digitization Areas and Technology:**

In reference to a previously indicated technological gap: an experienced equipment tech at Cal Poly, Doug Brewster, designed and manufactured a secure mechanism for mounting the imaging camera atop the lightbox. These camera mounts were assembled at Cal Poly and distributed to each institution with the necessary hardware.

The primary technological gap we now face is web-hosting the jpeg images. The iDigBio Image Ingestion Appliance has not functioned properly for several institutions, and iDigBio IT support has not been available due to being short-staffed. Our interim solution will be to have the PM regularly upload these institutions' images from a working Image Ingestion Appliance instance. The PM met with iDigBio IT and discussed a future, sustainable option for web-serving jpeg images involving Internet Archive, and we are strong advocates of this future direction.

There has been some indication that iDigBio servers may not continue to support our Symbiota instance (i.e., our data portal) past the duration of iDigBio's funding, which is of cause for some concern. We will continue to remain in discussion with iDigBio about the future of our data and consider alternative options.

### **Share and Identify Opportunities to Enhance Training Efforts:**

The PM visited all 22 institutions on the grant from December 2018 – February 2019. In these visits, the PM set up and installed equipment; discussed workflow and protocols; brainstormed ways to avoid bottlenecks and overcome inefficiencies; trained PIs, staff, volunteers, and students in digitization protocols, using equipment and software, managing data in the CCH2 portal, and other tasks; and addressed other institutional needs and requests.

The PM conducted check-in meetings with each institution in November 2018. December check-in meetings were not conducted because the PM visited most institutions in person, and January check-in meetings were not conducted due to the holiday break. February check-in meetings are scheduled for the second week of February.

Two well-attended webinars were conducted in October 2018 and November 2018. The first described how to navigate and edit data in the CCH2 portal and had 12 participants. The second described how to use the administrator control panel and had four participants. Recordings of these webinars are posted on the CAP TCN website ([capturingcaliforniasflowers.org](http://capturingcaliforniasflowers.org)). Relevant webinars from other projects (e.g., the Symbiota Working Group) are also posted on the Webinars page of the website, and collaborators are regularly encouraged to visit the site for these and other resources. Future webinars will address the topics of collection management tools, data cleaning, and georeferencing.

The website has been kept up-to-date with the latest digitization protocols and new training resources. The PM produced an imaging flowchart to help student technicians quickly remember critical steps, and she completed an [imaging protocol video](#) that walks the viewer through the imaging process. The PM is currently working on a comprehensive digitization manual that will contain a project overview, workflow, setup description, all protocols, and other necessary information.

The PM is regularly on call and quick to respond to communications from collaborators regarding best practices, equipment issues, and protocol questions. She also communicates to collaborators that she is willing to schedule and conduct virtual trainings as requested.

The lead PI and PM are planning a TCN community event to coincide with the annual meeting of the California Botanical Society in San Luis Obispo on April 6, 2019. This will be an ideal opportunity to communicate our experiences and discuss best practices moving forward. The lead PI and PM will also attend Botany 2019 in late July, where they will be able to meet and discuss pertinent topics with collaborators.

#### **Share and Identify Collaborations with other TCNs, Institutions, and Organizations:**

Collection data from San Francisco State University have been successfully imported into the CCH2 portal and are now being actively managed there. The PM and DM provide training and support to SFSU workers as necessary. Additional herbaria other than those listed in the grant will soon be solicited for data to host in the CCH2 portal.

The PM attended the iDigBio Phenology Deep Learning Workshop on January 17-18, 2019, where the topic of storing phenological data from machine learning applications was discussed. We are continuing to consider themes from this workshop and how they relate to coding phenological data from our specimen images. For example, we will consult developers of the Plant Phenology Ontology while determining the necessary vocabulary for codifying phenological data in the CCH2 portal in a standardized way.

The PM has identified many iDigBio-related resources for dissemination to CAP collaborators including webinars for georeferencing and data cleaning recorded by the Symbiota Working Group. We intend to schedule a community viewing time for one or more of these webinars as the topics become relevant to the activities of the TCN.

#### **Share and Identify Opportunities and Strategies for Sustainability:**

Project leadership is training administrators and technicians in using the CCH2 data portal to enable active collections management in perpetuity.

Each institution has identified a sustainable data storage solution for their specimen images that does not rely solely upon external hard drives or subscription services. As iDigBio IT support wanes, we are in discussion with iDigBio IT to find a long-term, sustainable option for web-serving jpeg images for viewing on the CCH2 portal.

#### **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). Pictures of most collections on the grant were posted as the PM conducted site visits. PIs, volunteers, and students are posting about the project as digitization activities accelerate.

The PM presented a poster at the Southern California Botanists Symposium on November 3<sup>rd</sup> and is currently drafting a scholarly article to submit to *Madroño*, the peer-reviewed journal published by the California Botanical Society.

The Central California Coast National Public Radio station, KCLU, featured digitization through the CAP TCN at UC Santa Barbara in this news piece: <http://www.kclu.org/post/south-coast-scientists-involved-creation-online-plant-encyclopedia-study-climate-change#stream/0>. An intern for the Cal Poly College of Science and Mathematics photo-documented the imaging training sessions at Cal Poly.

Cal Poly is tentatively planning an “herb-a-thon” community digitization event for local naturalists in the San Luis Obispo area during which we will discuss the importance of natural history collections for science and society, which an emphasis on phenological research.