

Mobilizing New England vascular plant data to track environmental change

Patrick W. Sweeney
Yale University Herbarium, Peabody Museum of
Natural History



Mobilizing New England Vascular Plant Specimen Data



to Track Environmental Changes

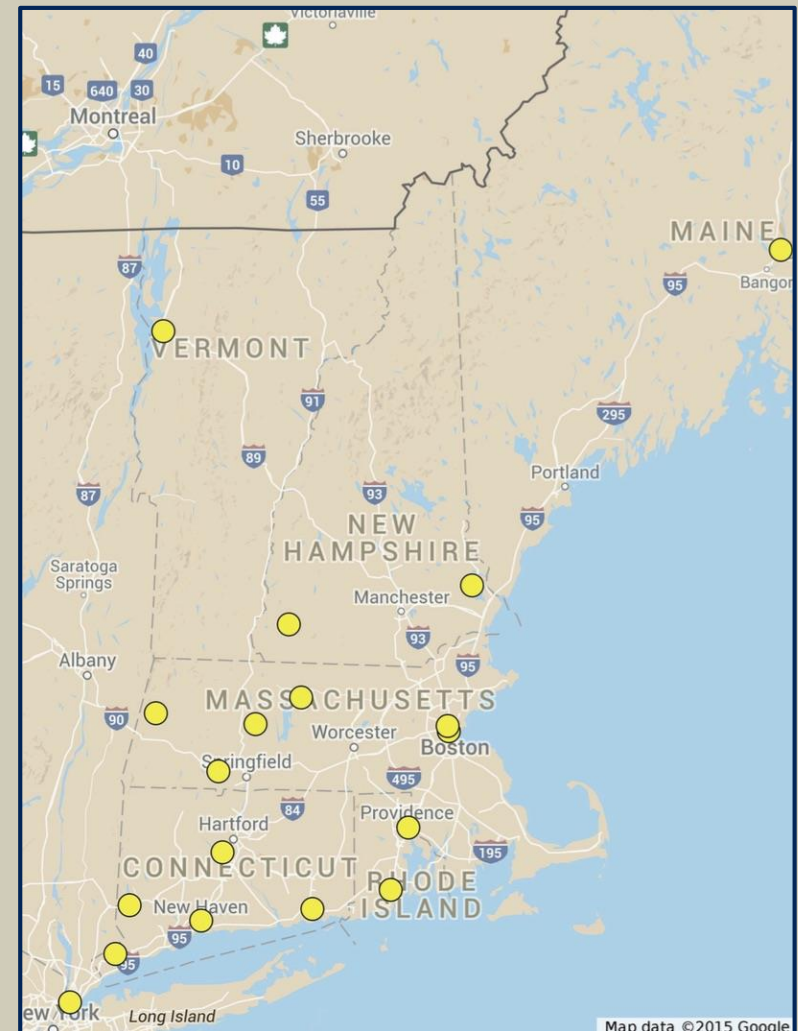


OUTLINE

- Management of the network including oversight and processes
- Data management & flow
- Open access of data
- Research use of data
- Status

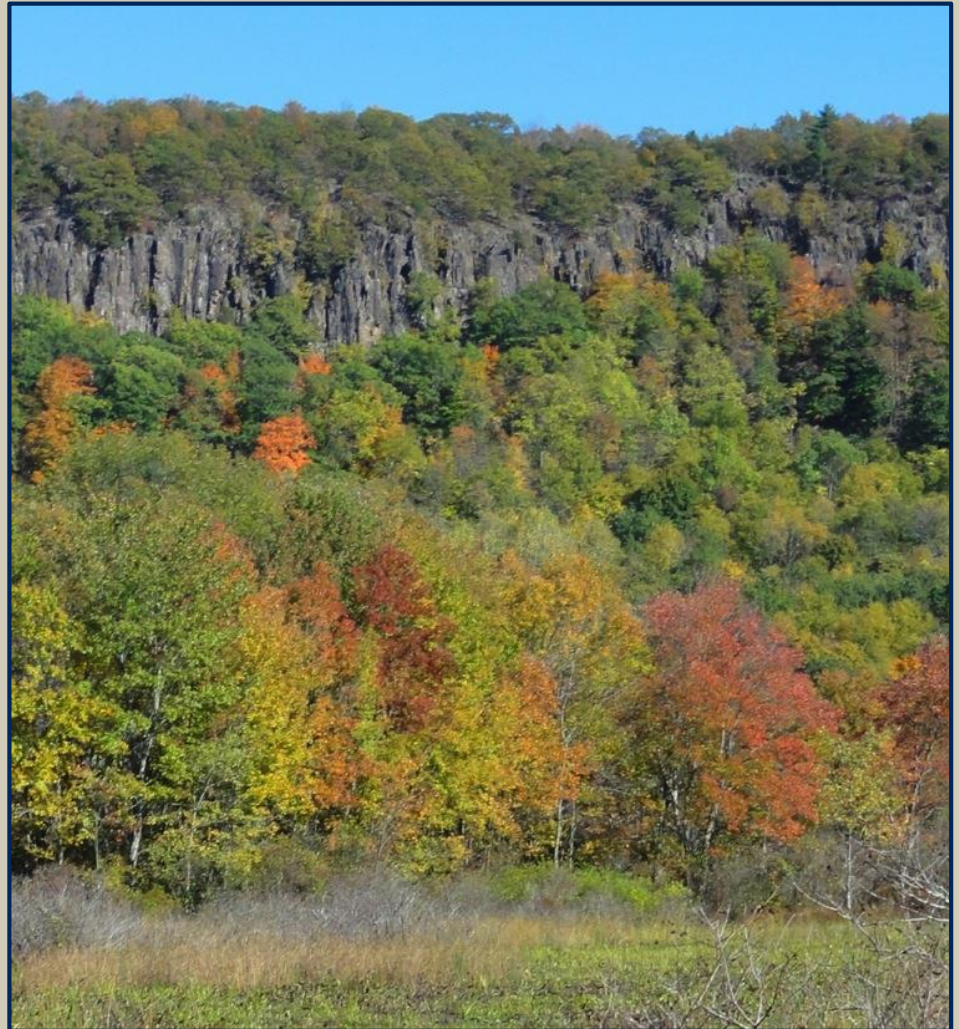
OVERALL OBJECTIVES

- Digitize 1.3 million New England vascular plant specimens from 15 regional herbaria



RATIONALE

- The main goal of this TCN is to provide data to support the study of the consequences of climate change and land use history in the New England region over the last two centuries



PARTICIPANTS

Digitizing Institutions:

- Brown University (BRU)
- Harvard University (HUH)
- U. of New Hampshire (NHA)
- U. of Massachusetts Amherst (MASS)
- U. of Vermont (VT)
- Yale University (YU)
- Connecticut College (CCNL)
- Harvard Forest (HF)
- Keene State (KESC)
- U. of Rhode Island (KIRI)
- Western Connecticut State U. (WCSU)
- Westfield State U. (WSCH)

Partners:

- Bartlett Arboretum (BART)
- Berkshire Museum (BERK)
- Boston University (BSN)
- Central Connecticut State U. (CCSU)
- U. of Maine (MAINE)
- New York Botanical Garden (NYBG)

PENs:

Engineering group:

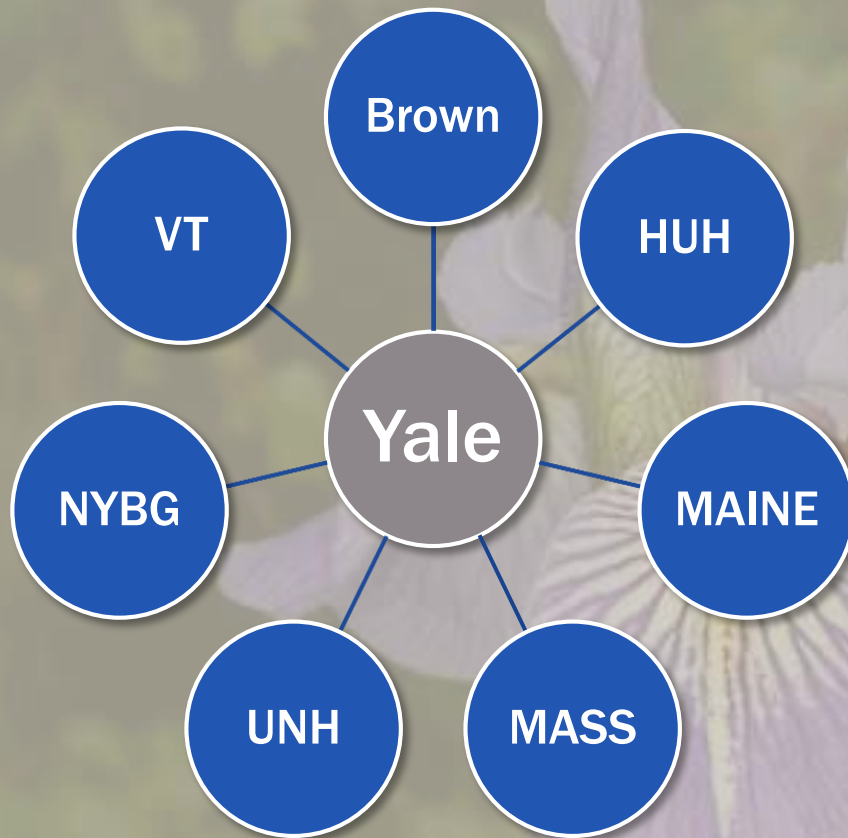
- North Carolina State U.

PROJECT MANAGEMENT

**TCN:
NEVP**

- **YU Overall Management**
- **Annual Meetings**
- **Email Communication**
- **Open Atrium – web-based collaboration tool**
- **Website**

PROJECT MANAGEMENT



■ Overall Mgmt.

- informatics & admin. support
- software, hardware, & workflow dev.
- subcontract with NCSU (conveyor system)
- dedicated point person
- progress tracking, reporting, etc.

■ Each institution autonomous

- develop and customize workflows
- digitize
- Informatics support
- initiate data and image transfer
- report to lead institution

PROJECT MANAGEMENT

**TCN:
NEVP**

- YU Overall Management
- Annual Meetings
- Email Communication
- **Open Atrium – web-based collaboration tool**
- Website

PROJECT MANAGEMENT



psweeney [Home](#) > [Groups](#) > [Digitization](#) 🔒 **NEVP**

Case Tracker

[+](#) Create content [⚙️](#) Settings [🔍](#) Search

[View](#) **Cases** [Edit](#) [Revisions](#) [Archive](#) [Projects](#) [Add Project](#) [Add Case](#) [?](#)

Priority	Last post	Title	Assigned	Status	Due Date
PD	1:38am May 18, 2015	New UI: Beginning Date field	yiming	Open	Apr 5, 2015
PD	4:55pm Apr 2, 2015	Metadata Editor: Beginning Date Field	yiming	Open	Apr 7, 2015
PD	4:16pm Mar 31, 2015	RunOnceAsAdmin	yiming	Resolved	Mar 21, 2015
PD	11:31am Mar 31, 2015	Collector Names	yiming	Resolved	Mar 18, 2015
PD	11:31am Mar 31, 2015	Unflag and Folder	yiming	Resolved	Mar 18, 2015
PD	2:17pm Mar 13, 2015	Visual Indicator for Mult. Spec. Sheets	yiming	Resolved	Jun 8, 2014
PD	2:15pm Mar 13, 2015	Flagging Issue	yiming	Resolved	Nov 26, 2014
PD	2:13pm Mar 13, 2015	UI data fields and tabbing	yiming	Resolved	Jun 17, 2014
PD	11:36am Mar 12, 2015	Shoot image retry option	yiming	Resolved	Dec 25, 2013

Filter results

Search

Priority

Status

Show closed cases

[Apply](#)

My cases

No cases found

<http://www.openatrium.com>

PROJECT MANAGEMENT

Mobilizing New England Vascular Plant Specimen Data



to Track Environmental Changes

HOME

PROJECT SUMMARY

PARTICIPANTS

PORTAL

PROJECT MANAGEMENT

RESOURCES

Resources

The digitization plan for the NEVP TCN has three main components:

1) **Pre-capture.** In this stage data will be captured that reflects the physical storage structure of the collection before imaging and databasing individual specimens, and then associating this pre-captured data with specimen occurrence records.

2) **Primary Digitization.** In the primary digitization stage, an image, a barcode number, and a subset of label data will be captured and pre-capture data will be associated with specimen occurrence records.

<http://nevp.org>

PROJECT MANAGEMENT

Consortium of Northeastern Herbaria

- **Maintains Portal**
 - **informatics Support**
- **Governing Structure**
- **Annual Meetings**
- **Drupal Website**
 - **communications, business, etc.**

PROJECT MANAGEMENT



[About](#) | [Portal](#) | [Membership](#) | [Governance](#) | [Meetings](#) | [Resources](#)

Northeastern Herbaria

- ▶ [Canada](#)
- ▶ [United States](#)

[Join CNH!](#)

[Other Nodes](#)

Governance

Interim assignments for governance of the CHN were established among the attendees of our 2008 meeting. A more formal governing structure was adopted at our 2009 meeting ([download](#)).

Steering committee

The committee includes individuals focusing on particular tasks as well as those with special expertise. The current members of the committee are:

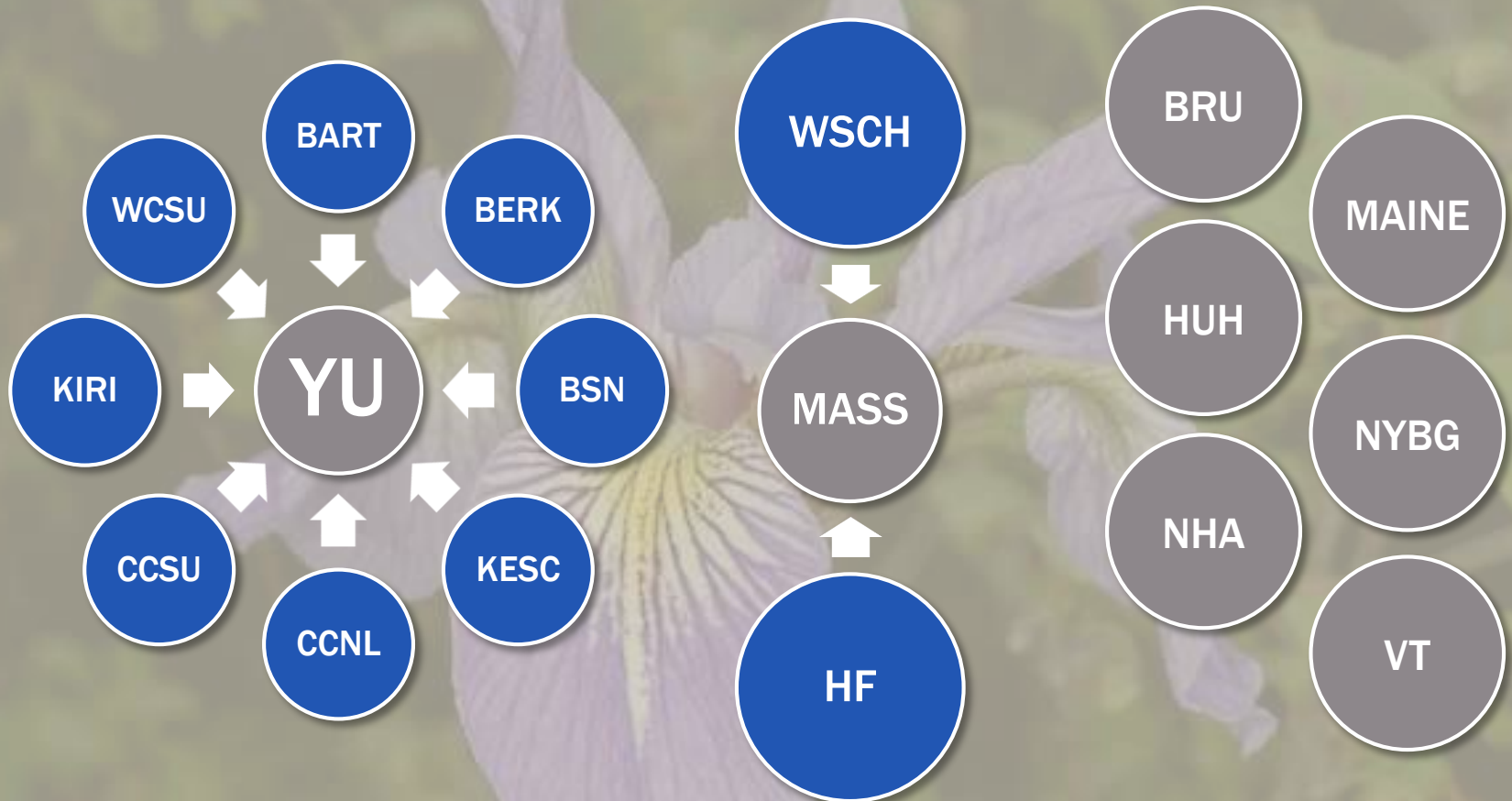
Patrick Sweeney, Yale University, chairman
Dorothy Allard, University of Vermont
Bob Capers, University of Connecticut, recording secretary

Search this site:

Search

<http://neherbaria.org>

DIGITIZATION: ORGANIZATION



DIGITIZATION: WORKFLOW

Pre-capture



Primary
digitization



Secondary
Digitization (data
enhancement)

WORKFLOW: PRIMARY DIGITIZATION

1° Digitization Systems

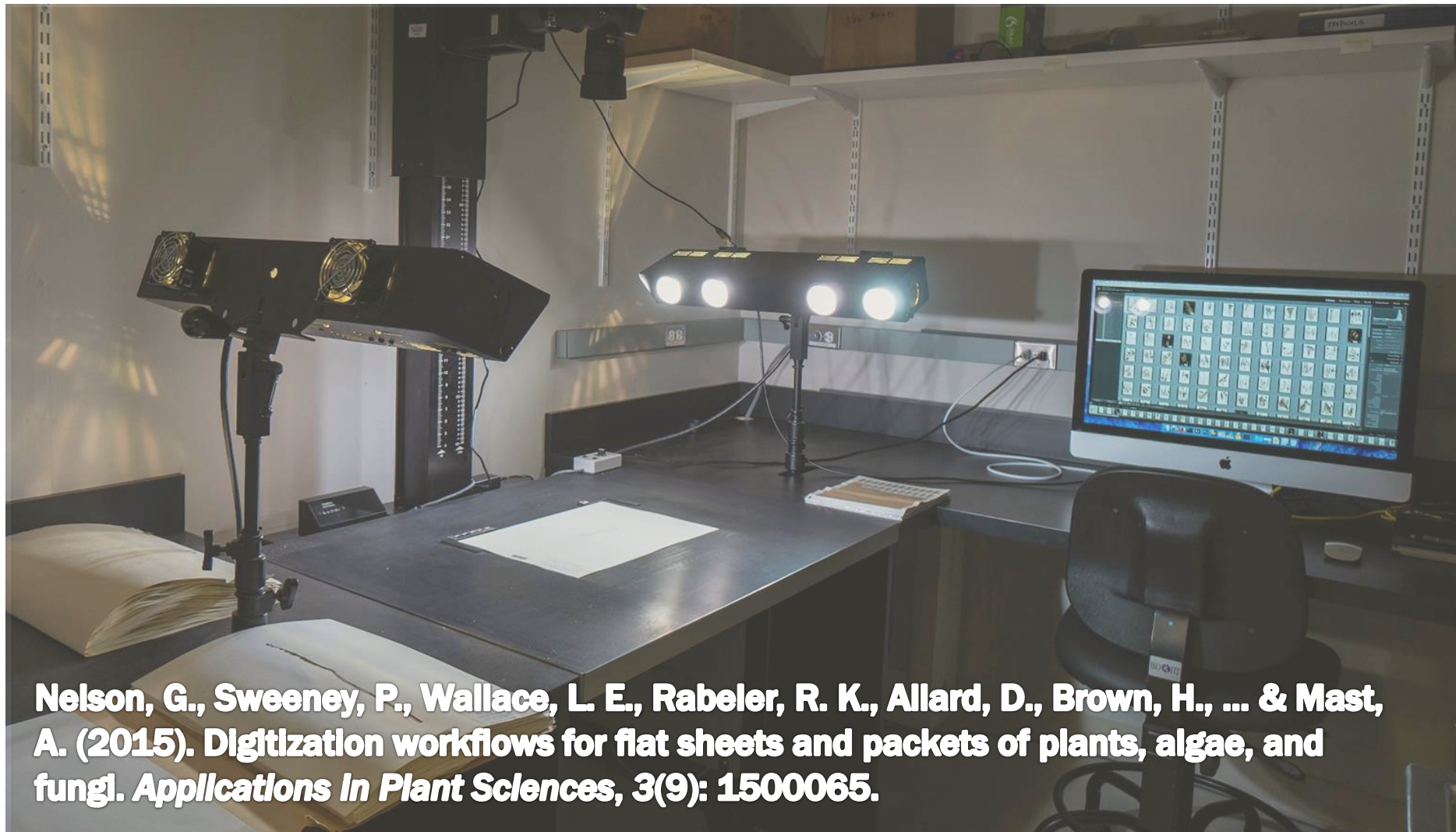
Light Stand/Box

Conveyor

WORKFLOW: PRIMARY DIGITIZATION - CONVEYOR

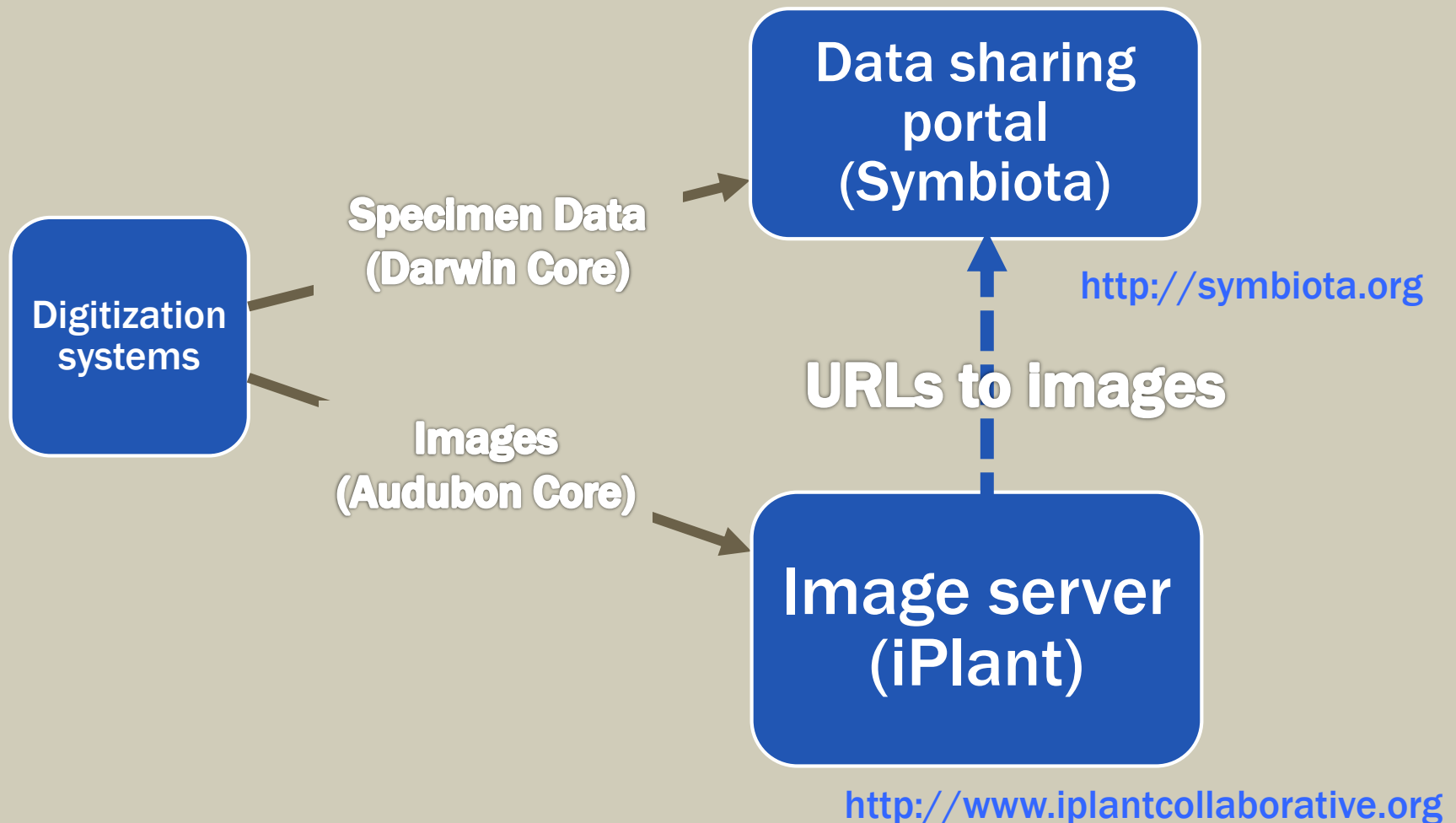


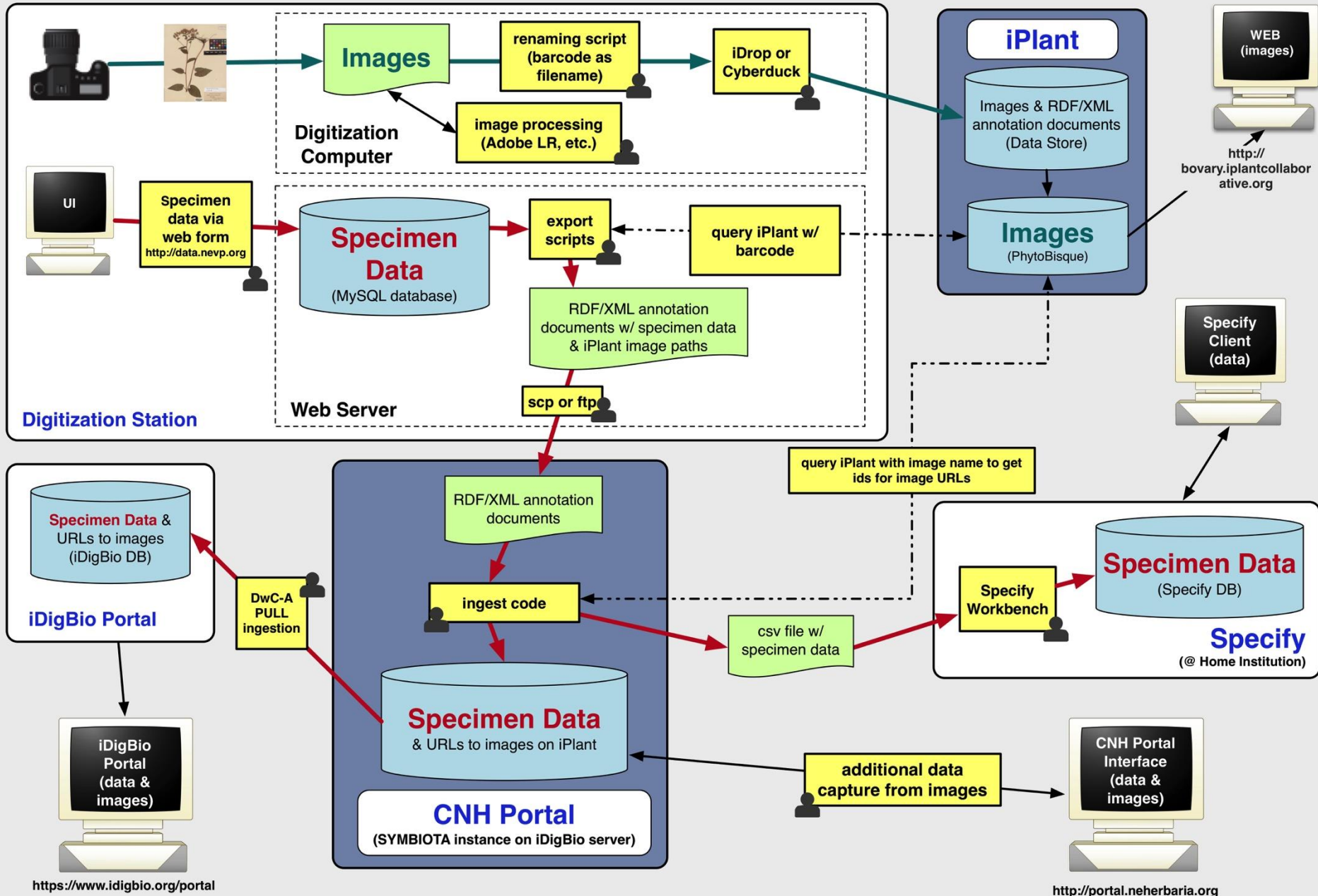
WORKFLOW: PRIMARY DIGITIZATION - LIGHT STAND



Nelson, G., Sweeney, P., Wallace, L. E., Rabeler, R. K., Allard, D., Brown, H., ... & Mast, A. (2015). Digitization workflows for flat sheets and packets of plants, algae, and fungi. *Applications In Plant Sciences*, 3(9): 1500065.

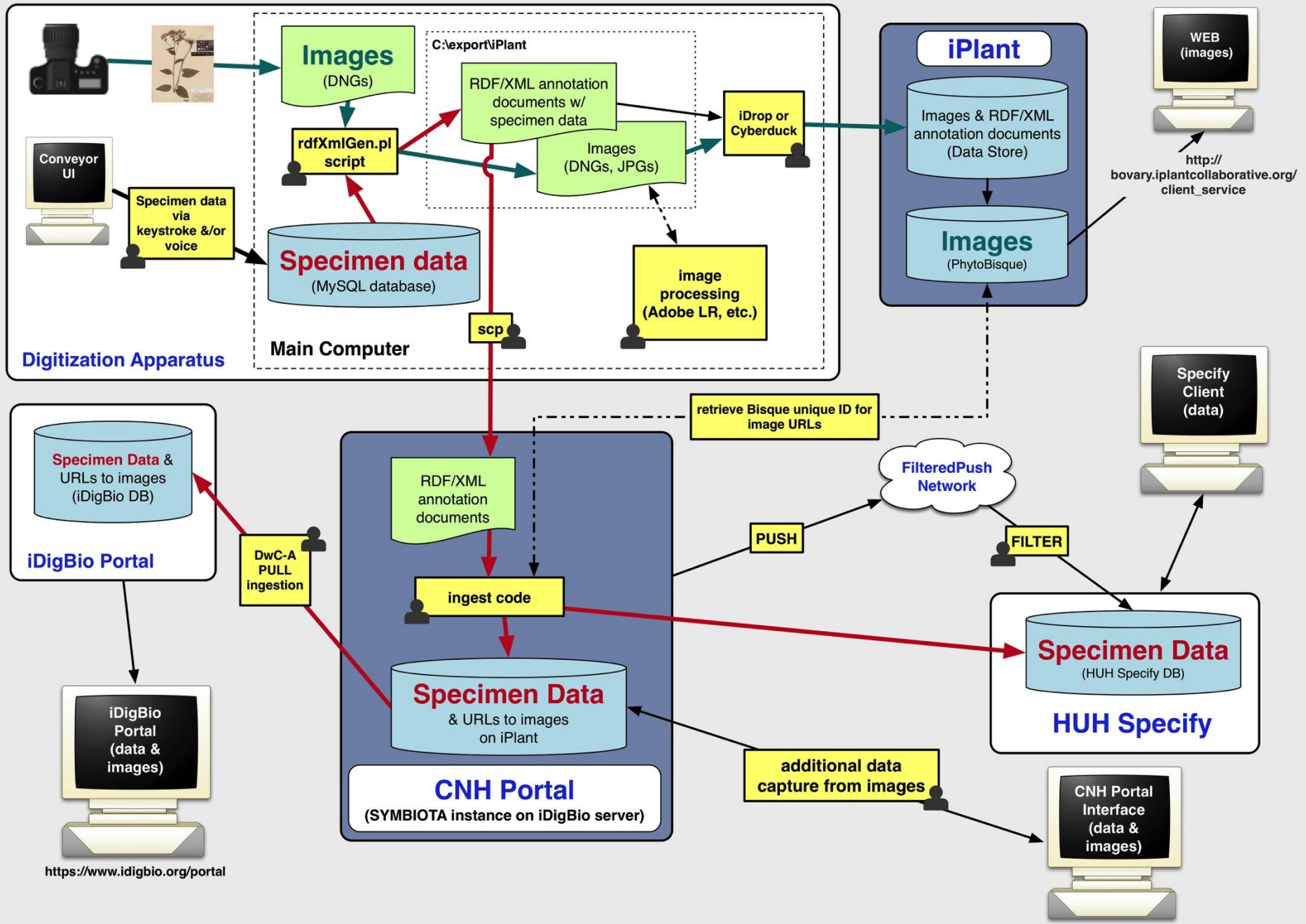
WORKFLOW: PRIMARY DIGITIZATION, DATA FLOW & MANAGEMENT





NEVP data flow diagram, web GUI & light stand/box, simplified (2015-10-23, vers. 1.3)


process
data
storage
 = person involved



NEVP data flow diagram, HUH/conveyor simplified (2015-10-23, vers. 5) process data storage = person involved

ACCESS: PORTAL

- All of the data and images are available to public through *CNH* portal
- Data capture and editing functionality
- Publish to iDigBio



[About](#) | [Portal](#) | [Membership](#) | [Governance](#) | [Meetings](#) | [Resources](#)

Portal Menu
Text-based Search
Map-guided Search
Collections
Species Lists
Sitemap

[Log In](#)

[New Account](#)

Herbarium Specimen Data Sharing Portal for CNH

Number of records in database: 906173

About:

The CNH portal provides access to herbarium specimen data housed in member institutions, with particular emphasis on specimens collected in the region. The database includes taxa traditionally found in herbaria, including plants, fungi, diatoms, algae, and lichens.

Use of any specimen data and related material (e.g., images, species checklists, etc.) accessed through this portal requires agreement to the terms and conditions in the [CNH data usage policy](#).

If your institution is interested in sharing data and is willing to abide by the terms of our [data sharing and data usage policies](#), email [Patrick Sweeney](#) for further instructions about how to make this happen.

Acknowledgements:

The CNH specimen portal utilizes the Symbiota framework. The Symbiota Software Project (www.symbiota.org) is an NSF funded endeavor based at Arizona State University. We are particularly indebted to Edward Gilbert for assisting CNH in implementing this Symbiota instance.

<http://portal.neherbaria.org>

RESEARCH USE OF DATA

- Global Change Studies (phenology data)

- R. Primack et al., fruiting phenology (in prep.)
- C. Davis et al., invasive species phenology (in prep.)
- Bellemare, J., & Deeg, C. (2015). Horticultural Escape and Naturalization of *Magnolia tripetala* in Western Massachusetts: Biogeographic Context and Possible Relationship to Recent Climate Change. *Rhodora*.

- Floristics

Angelo, R. and D.E. Boufford. 2014. Atlas of the flora of New England: Families of vols. 9-13 and 15-15 (except Fabaceae & Rosaceae) of the Flora of North America series. *Phytoneuron* 2014(98), 1-60.

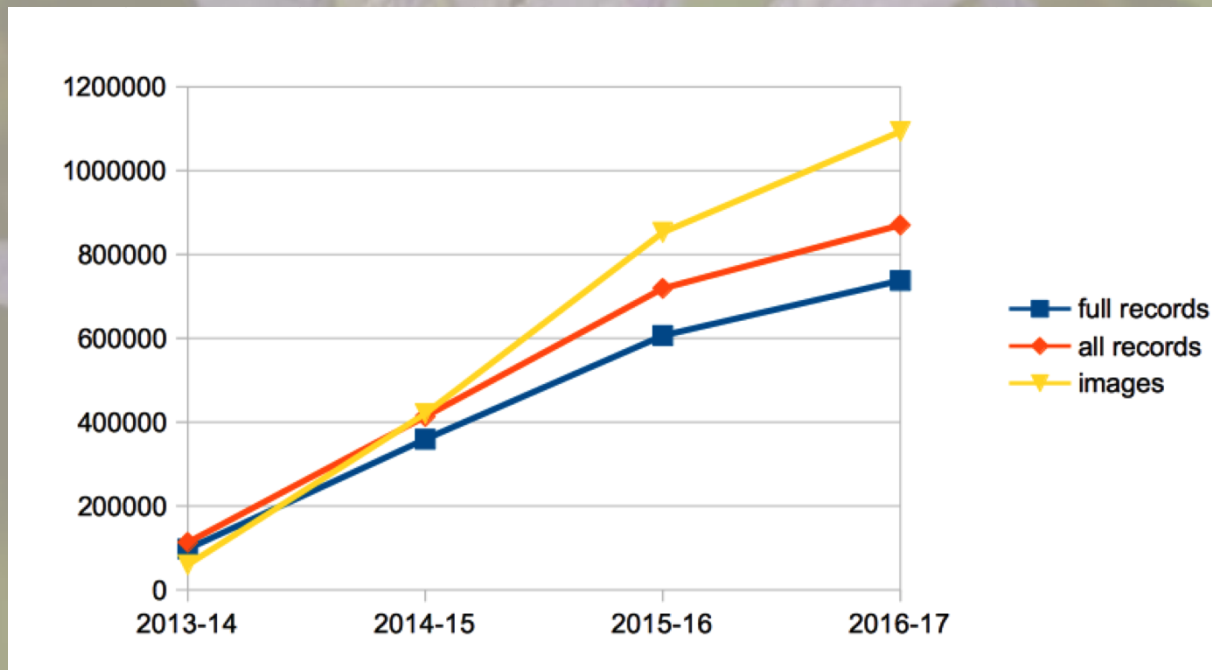
- Conservation/Management

- Education & Outreach

PROGRESS

■ Progress:

- Pre-capture – 800,000+ specimens
- Primary – ca. 400,000 over last year (over 600,000 to date)
- Secondary – New England gazetteer, phenology vocabulary, scoring underway




WORKFLOW: SECONDARY DIGITIZATION

- Capture phenology and habitat data from images of specimens and labels.

Occurrence Data	Determination History	Images	Admin
Collector Info			
atalog Number ?	Occurrence ID ?	Collector	Number Date
063041		William Russell Dudley	1904-07-14 <input type="button" value="Dupes"/>
ociated Collectors		Other Catalog Numbers ?	
Best Identification			
Scientific Name:		Author:	
Syringia divaricata		(L.) G.L. Nesom	
Qualifier: ?		Family:	Asteraceae
Identified By:		Date Identified:	
Locality			
Country	State/Province	County	Municipality
USA	Connecticut	New Haven County	
Locality:			
Locality Security			

Label Processing



ACKNOWLEDGEMENTS



National Science Foundation (EF1208829, EF1208835, EF1208972, EF1208973, EF1208975, EF1208989, EF1209149).



Symbiota Project



FilteredPush



iPlant Collaborative™ *Empowering A New Plant Biology*



Biota of North America

BONAP

