

# Mobilizing New England vascular plant data to track environmental change

P.W. Sweeney, D. Allard, D.S. Barrington, C.C.  
Davis, M.J. Donoghue, E.J. Edwards, D.R. Foster,  
P.J. Morris, C. Neefus, R.B. Primack, K.B. Searcy,  
B. Starly, J.R. Sullivan



Mobilizing New England Vascular Plant Specimen Data

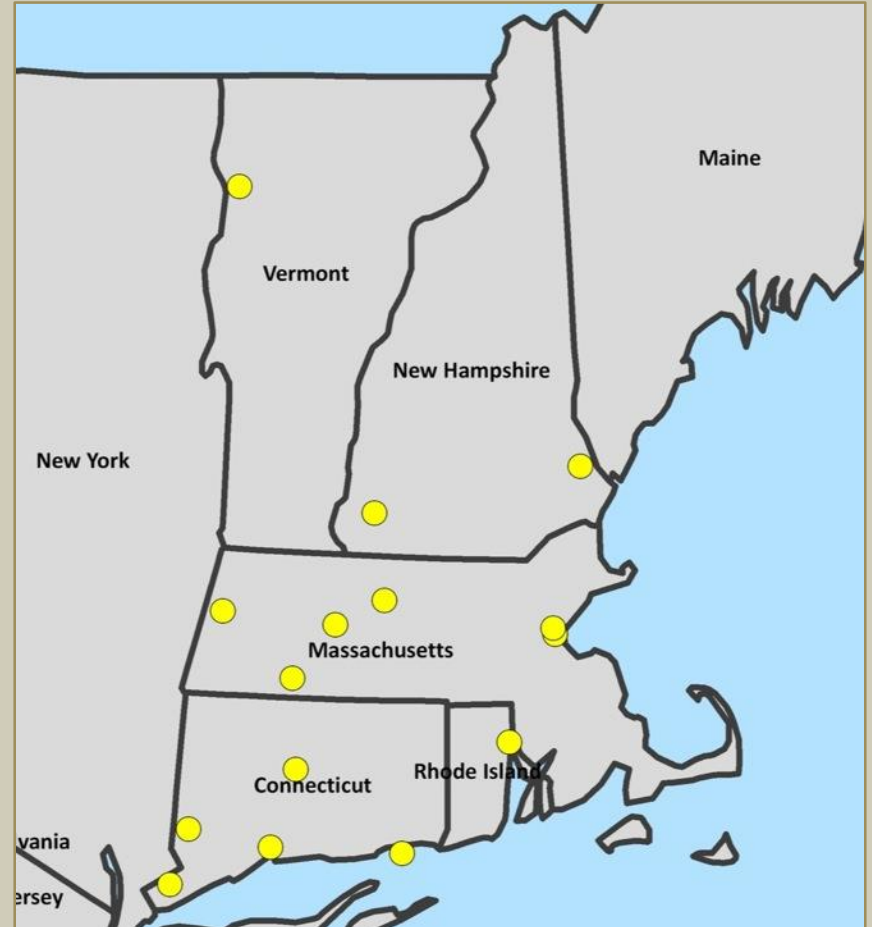


to Track Environmental Changes



# OVERALL OBJECTIVES

- Digitize 1.3 million N.E. vascular plant specimens from 15 regional herbaria



# RATIONALE

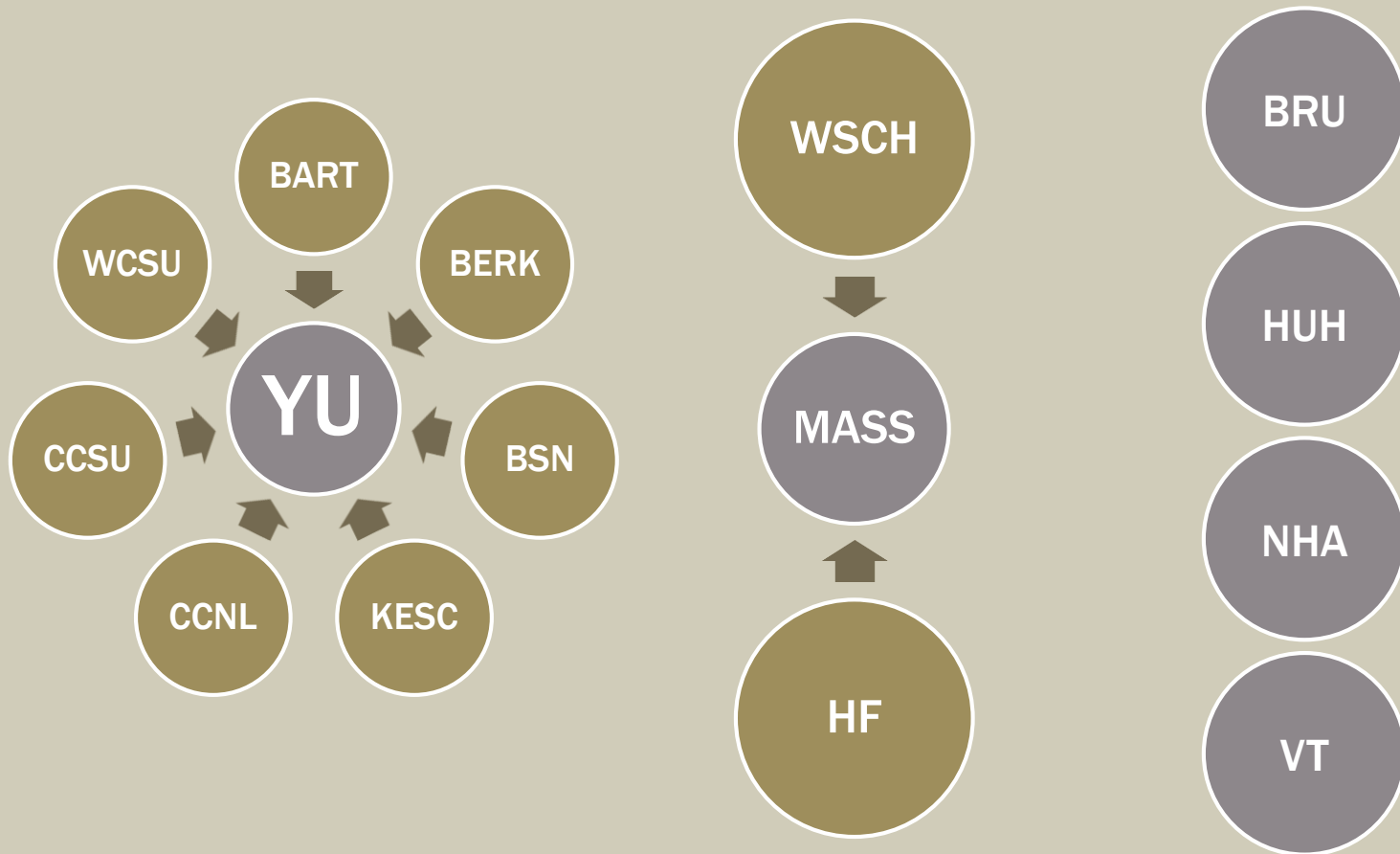
- The main goal of this TCN is to provide data to support the study of the consequences of climate change and land use history in the New England region over the last two centuries



# PARTNERS

- Brown University (BRU)
  - Harvard University (HUH)
  - U. of New Hampshire (NHA)
  - U. of Massachusetts Amherst (MASS)
  - U. of Vermont (VT)
  - Yale University (YU)
  - Boston University (BSN)
  - Bartlett Arboretum (BART)
  - Berkshire Museum (BERK)
  - Central Connecticut State U. (CCSU)
  - Connecticut College (CCNL)
  - Harvard Forest (HF)
  - Keene State (KESC)
  - Western Connecticut State U. (WCSU)
  - Westfield State U. (WSCH)
- Engineering group:
- North Carolina State U.

# DIGITIZATION PLAN: ORGANIZATION



# DIGITIZATION PLAN: WORKFLOW

Collection  
Preparation (Pre-  
capture)



Primary  
digitization



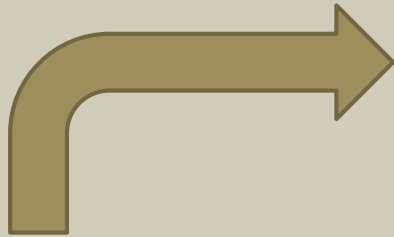
Data  
Enhancement  
(Secondary  
Digitization)

# WORKFLOW: PRE-CAPTURE

- Capture data (taxon, state) that reflects the physical storage structure of the collection before imaging and data basing individual specimens
- Associate this pre-captured data with specimen records at a later stage



# WORKFLOW: PRE-CAPTURE



*Cardamine maxima* (Nutt.) Alph. Wood  
Massachusetts  
YU



File Window Help

Folders Inventory

Filter Taxa by: Family  Genus

Genus	Specific E...	Infraspecif...	Infraspeci...	Scientific...	Authorsh
Cardamine	concatenata				(Michx.) Sw...
Cardamine	maxima				(Nutt.) Alph. W
Cardamine	diphylla				(Michx.) Alph. 1

Taxon

Genus

Specific Epithet

Infraspecific Rank

Infraspecific Epithet

Scientific Name

Authorship

State

County

Collector

Storage Location

Herbarium Acronym

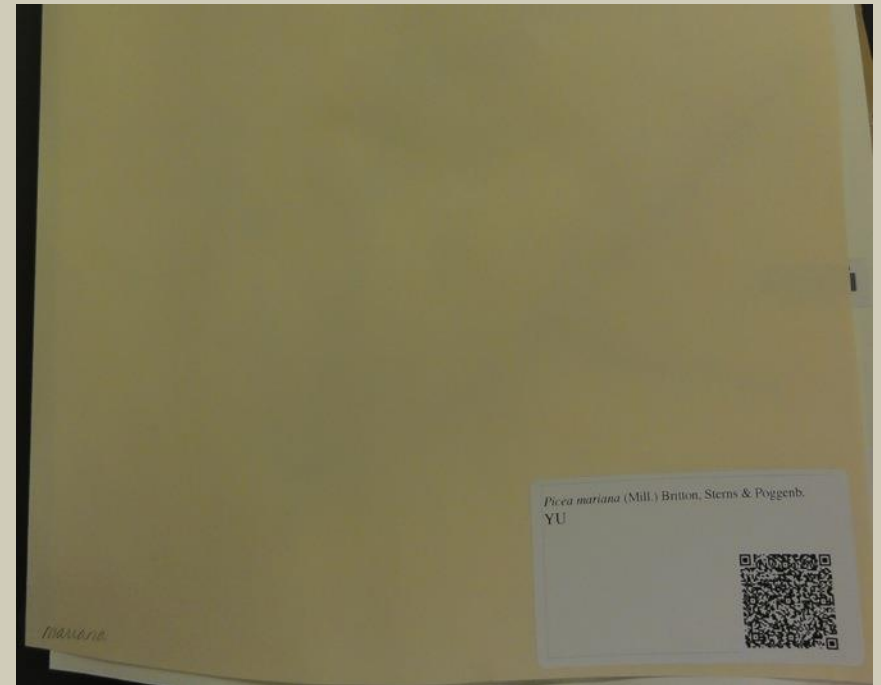
Number To Print

Print One Add to Print List

Print Clear

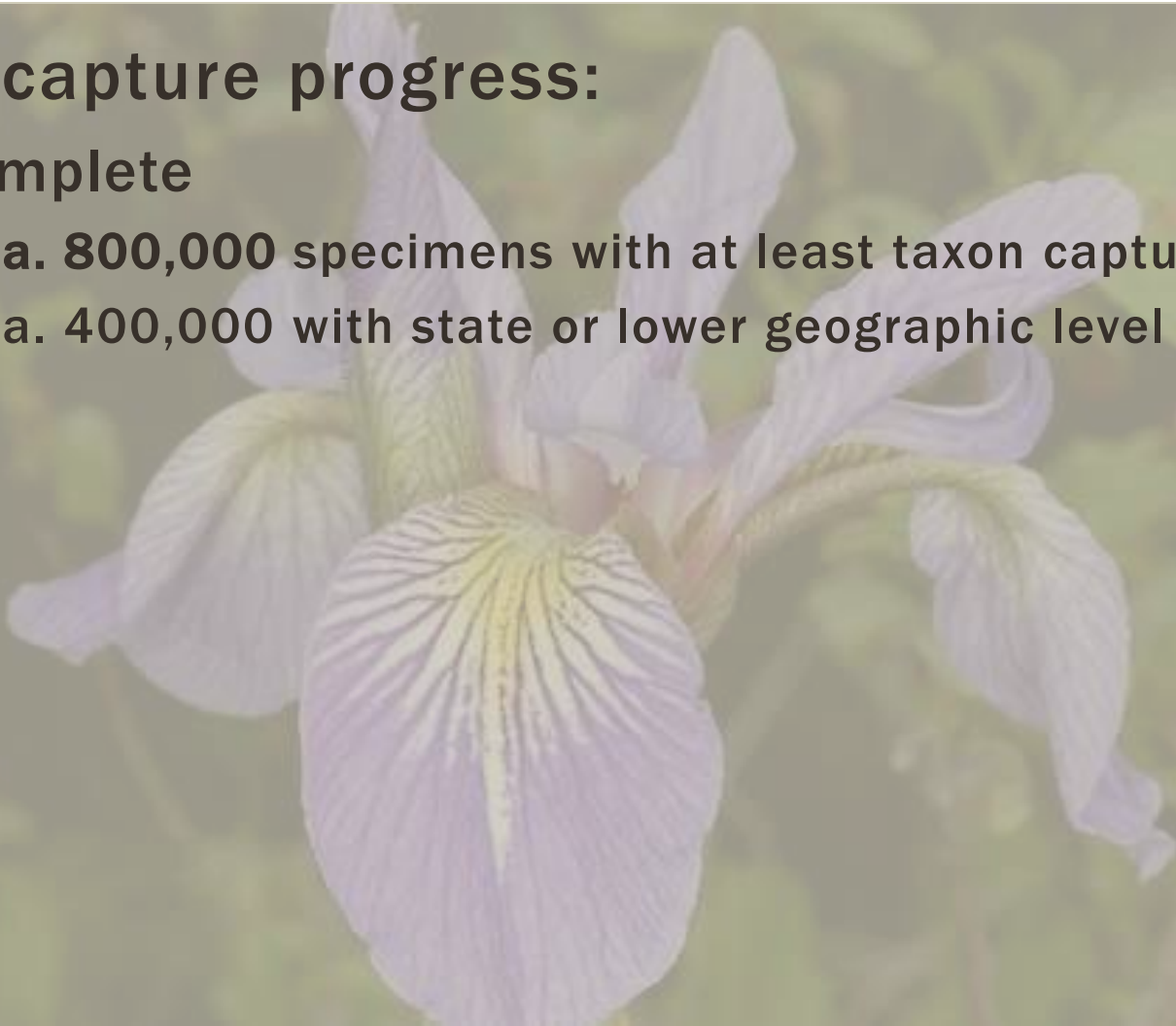


# WORKFLOW: PRE-CAPTURE



# WORKFLOW: PRE-CAPTURE

- **Pre-capture progress:**
  - **Complete**
    - **ca. 800,000 specimens with at least taxon captured**
    - **ca. 400,000 with state or lower geographic level**



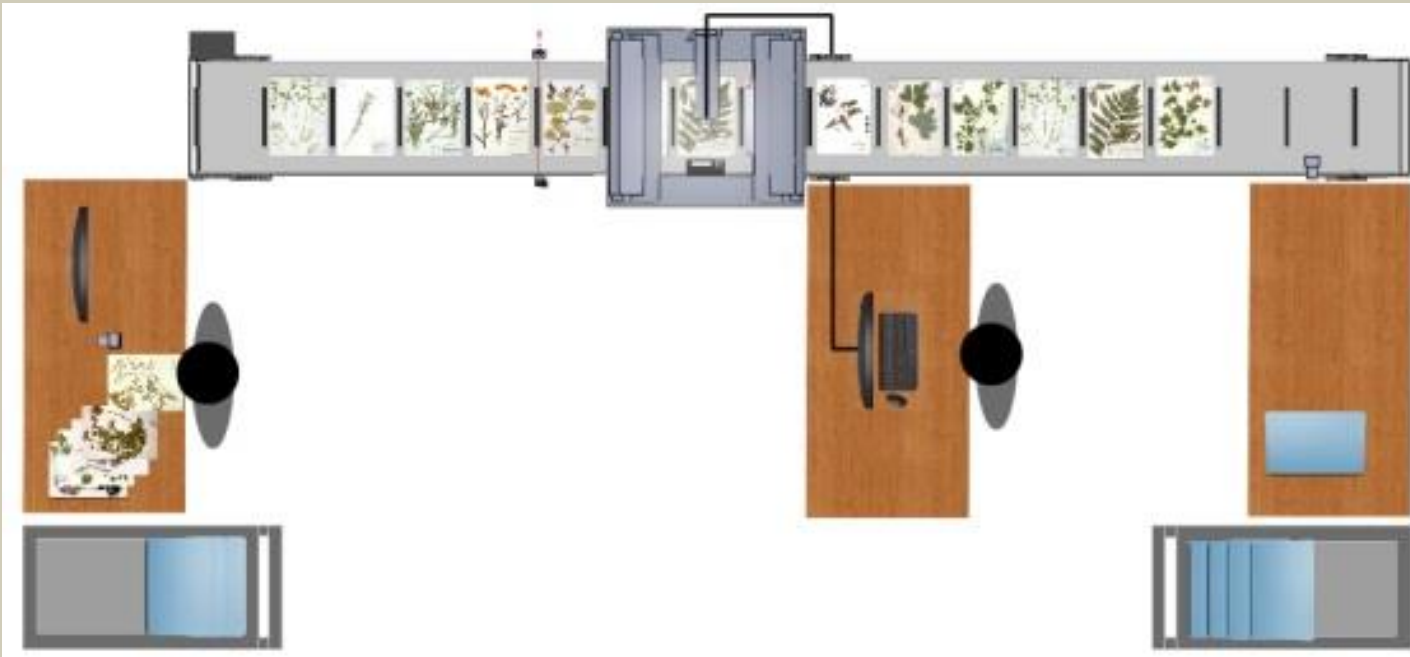
# WORKFLOW: PRIMARY DIGITIZATION

- Capture an image, a barcode number, a subset of label data & associate pre-capture data with specimen occurrence records

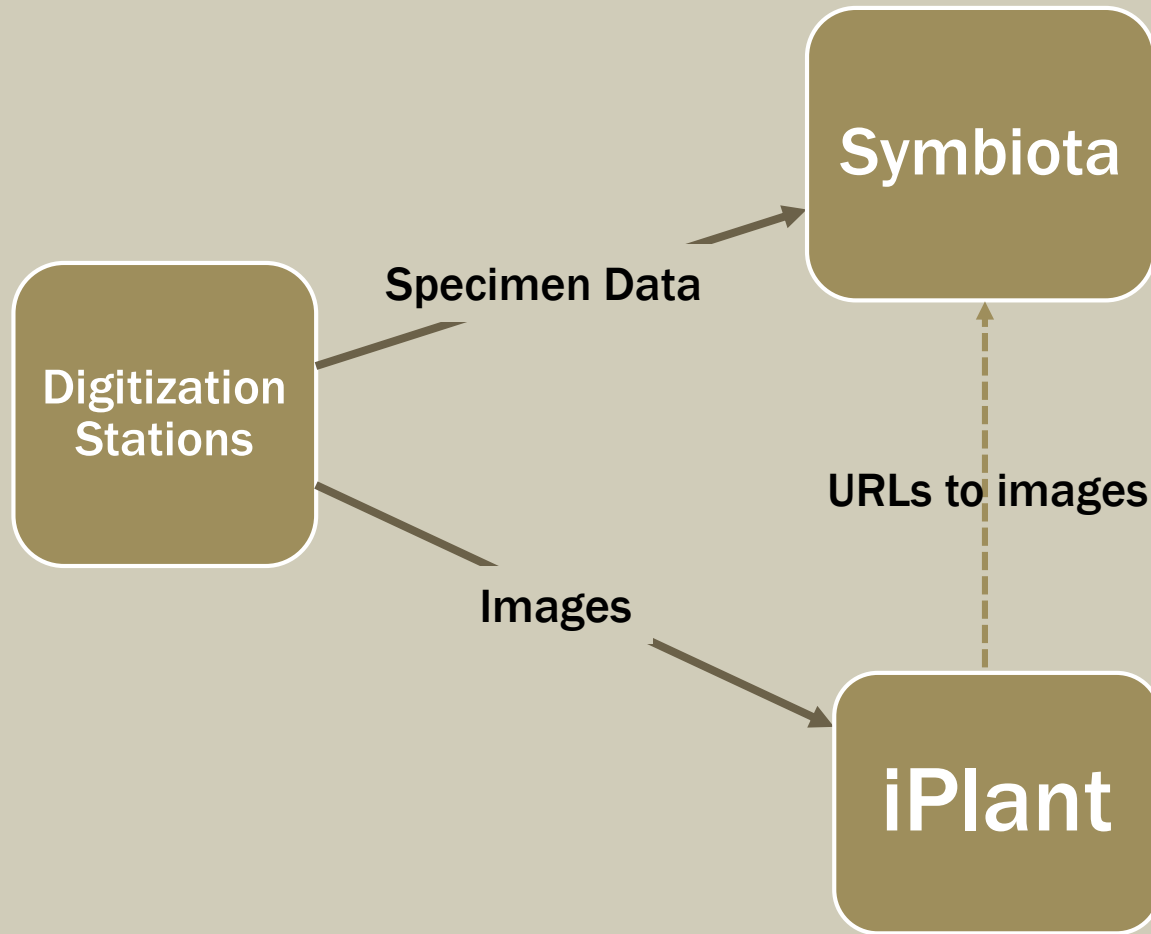


# WORKFLOW: 1° DIGITIZATION - CONVEYOR

- To increase the efficiency of capturing an image and specimen-level data, we are developing a high throughput digitization apparatus



# WORKFLOW: 1° DIGITIZATION



# WORKFLOW: PRIMARY DIGITIZATION

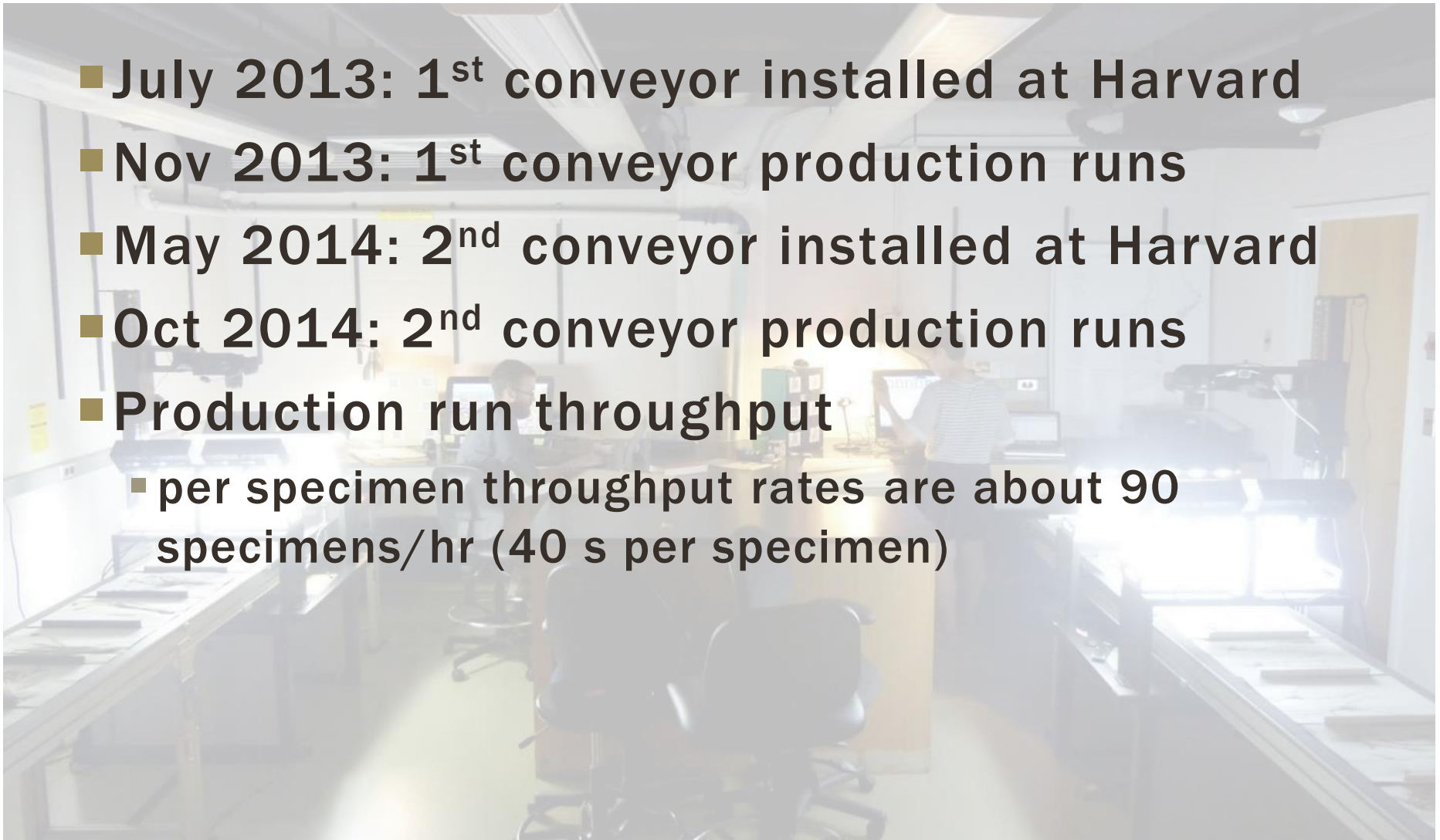
1° Digitization

Light box

Conveyor

# WORKFLOW: 1° DIGITIZATION - CONVEYOR

- July 2013: 1<sup>st</sup> conveyor installed at Harvard
- Nov 2013: 1<sup>st</sup> conveyor production runs
- May 2014: 2<sup>nd</sup> conveyor installed at Harvard
- Oct 2014: 2<sup>nd</sup> conveyor production runs
- Production run throughput
  - per specimen throughput rates are about 90 specimens/hr (40 s per specimen)



# WORKFLOW: 1° DIGITIZATION - CONVEYOR





# WORKFLOW: 1° DIGITIZATION - CONVEYOR

Herbarium Digitalization by OU

Settings About Help Logout

006107 89

006107 88

Camera Position

Scientific Name

Barcode

Genus

Collector

Species

Collector Number

Infraspecific Rank

Verbatim Collection Date

Infraspecific Epithet

Beginning Date

Scientific Name Author

Ending Date

Collection Code

Town

Identification Qualifier

County

Record Entered By

State

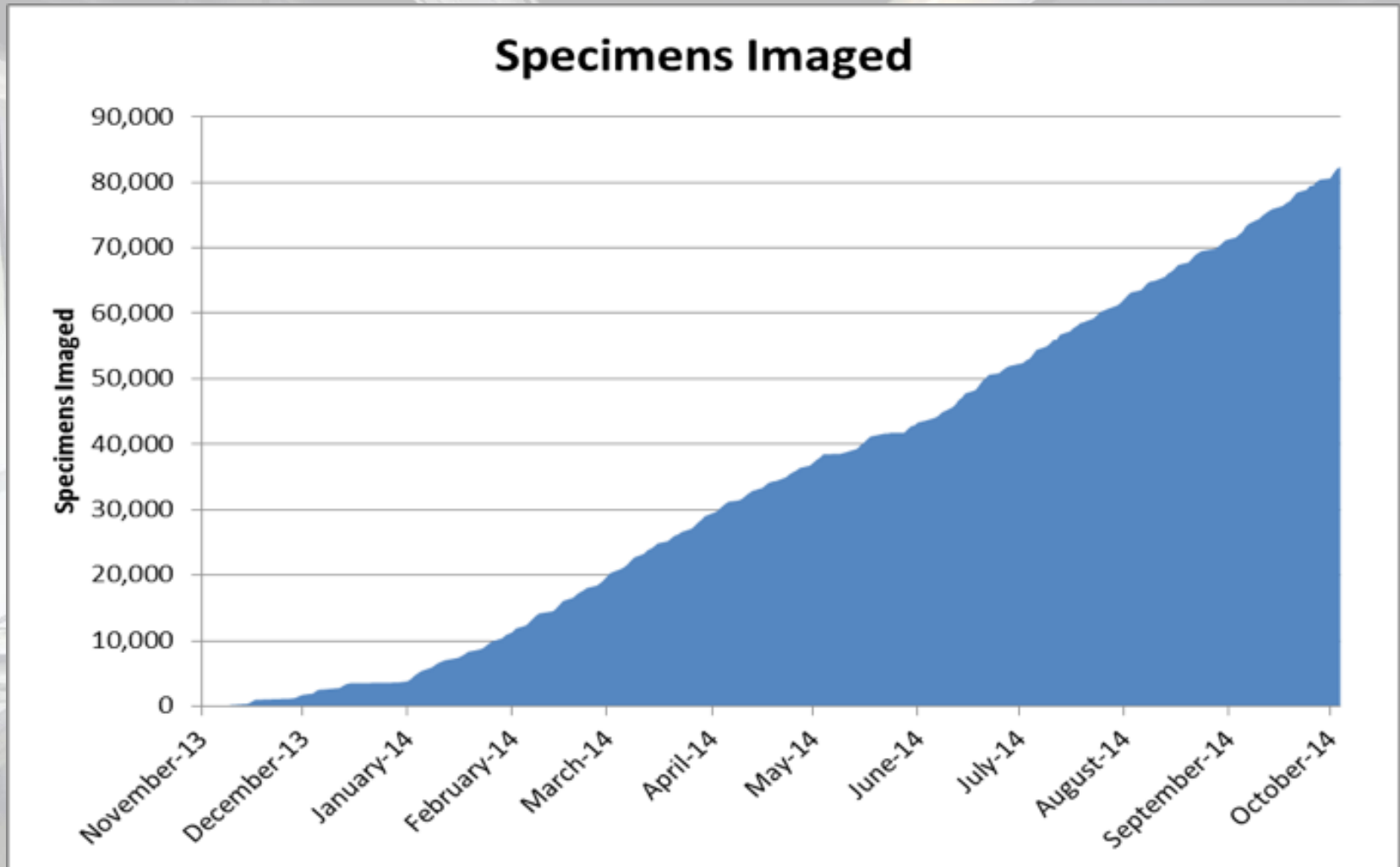
Tags

- Flag Specimen
- Folder

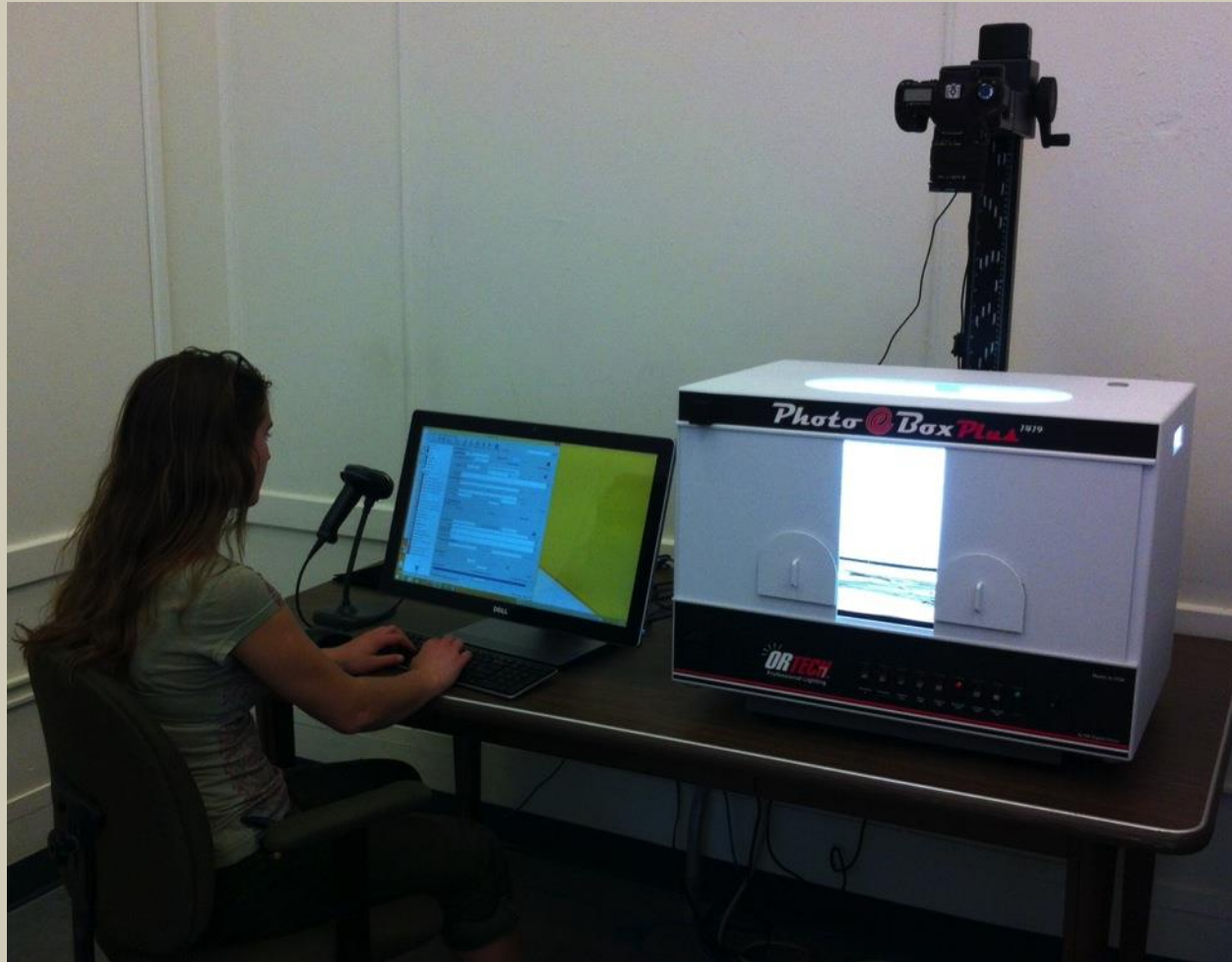
66 Connected

9:22 AM 6/18/2014

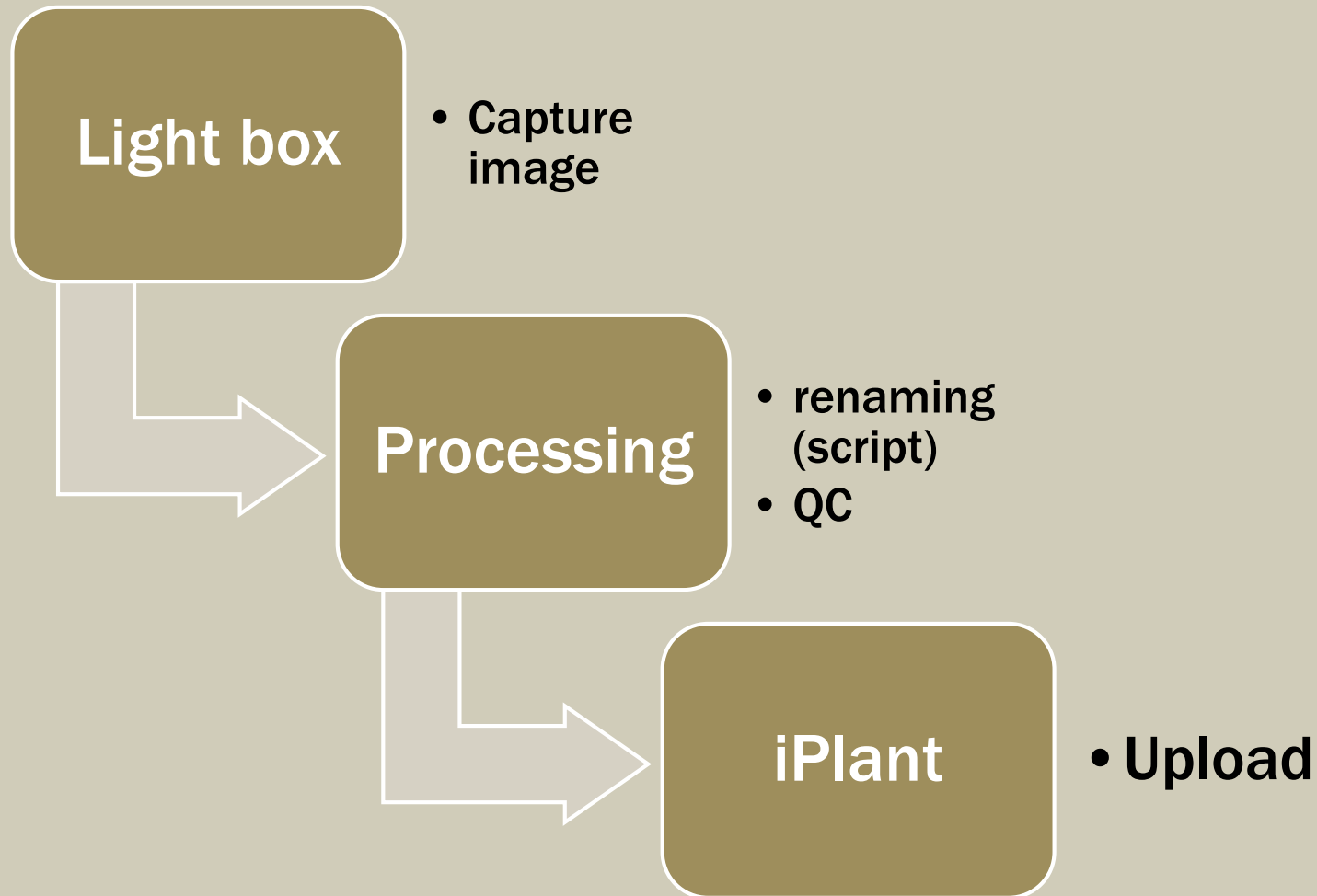
# WORKFLOW: 1° DIGITIZATION - CONVEYOR



# WORKFLOW: 1° DIGITIZATION - LIGHT BOX



# WORKFLOW: 1° DIGITIZATION - LIGHT BOX - **IMAGING**



# WORKFLOW: 1° DIGITIZATION - LIGHT BOX - **DATA CAPTURE**

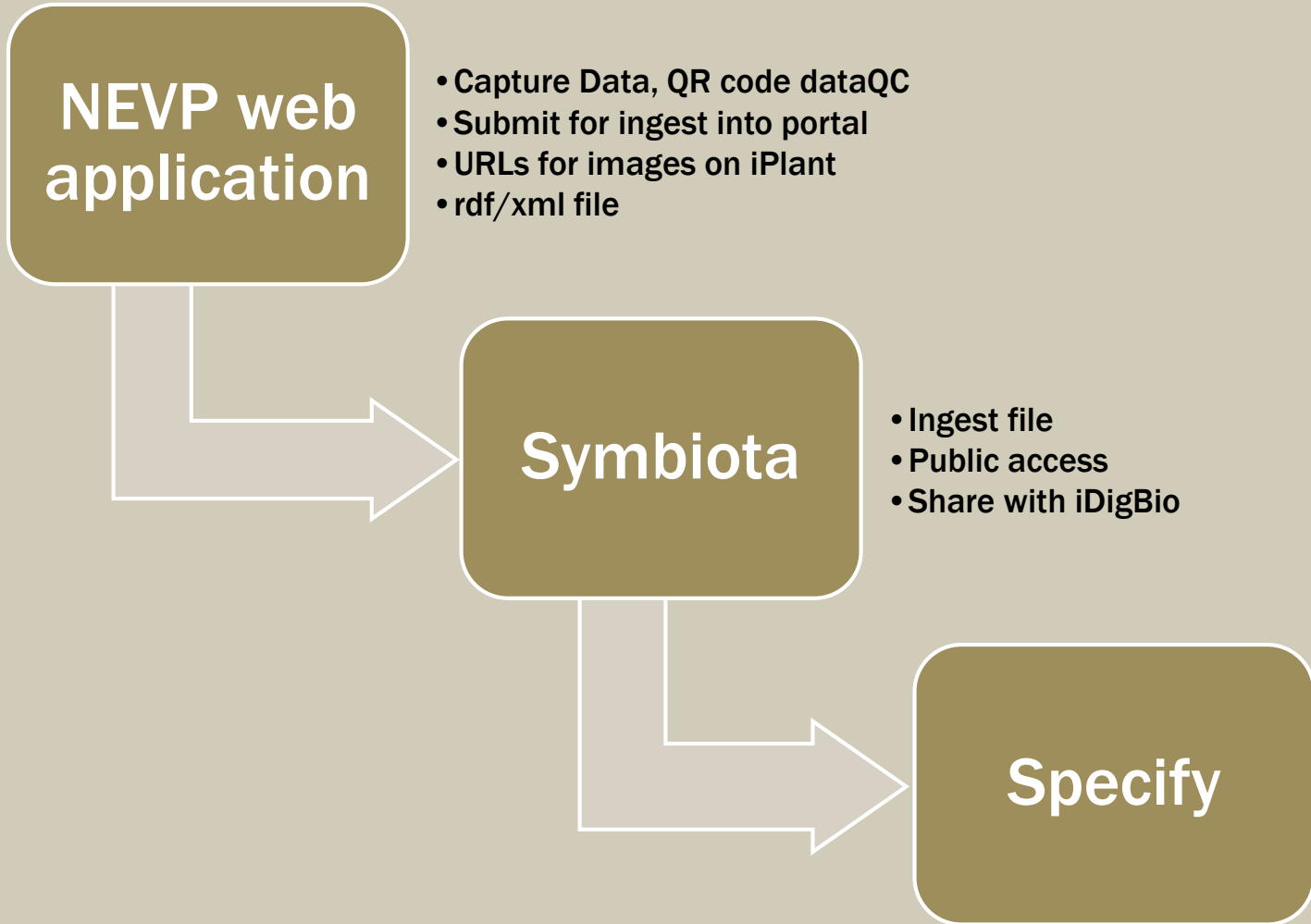
**NEVP web  
application**

- Capture Data, QR code dataQC
- Submit for ingest into portal
- URLs for images on iPlant
- rdf/xml file

**Symbiota**

- Ingest file
- Public access
- Share with iDigBio

**Specify**



# WORKFLOW: 1° DIGITIZATION - LIGHT BOX - DATA CAPTURE

## NEVP Specimen Data Entry Form

You are logged in as: [Patrick Sweeney](#) [LOGOUT](#) | [SPECIMEN TABLE](#) | [COLLECTOR TABLE](#)

Don't forget to logout when your session is complete. If you are not the user listed above, please logout.

QR Code:   Reuse QR Code  
Barcode:  Collection Code:   Reuse Collection Code

### SCIENTIFIC NAME DATA

Scientific Name:   Reuse taxon

[ENTER NEW TAXON](#)

[HELP](#)

### COLLECTOR DATA

Collector:  Coll Number:   
Coll Date (Earliest):  (Latest):  Verbatim Coll Date:

[HELP](#)

### LOCALITY DATA

Country:  State/Province:  County:   
City:  Locality:

User:

Date Recorded:

# WORKFLOW: 1° DIGITIZATION - LIGHT BOX - DATA CAPTURE

<a href="#">YU.074948</a>	<i>Epilobium ciliatum</i> Raf.	C. W. Swan			1882-09-01	Massachusetts		Tewksbury		YU
<a href="#">YU.074947</a>	<i>Epilobium ciliatum</i> Raf.	C. W. Swan		6-21-81	1881-06-27	Massachusetts		Boston	W Boston flats	YU
<a href="#">YU.074959</a>	<i>Epilobium coloratum</i> Biehler	Swan, C. W.			1885-08-20	Massachusetts		Westford	Graniteville	YU
<a href="#">YU.074967</a>	<i>Epilobium hirsutum</i> L.	C. W. Swan		6.IX.99	1899-09-06	Massachusetts				YU
<a href="#">YU.074985</a>	<i>Epilobium leptophyllum</i> Raf.	C. W. Swan		8_22_84	1884-08-22	Massachusetts		Bedford		YU
<a href="#">YU.074997</a>	<i>Epilobium palustre</i> L.	C. W. Swan			1882-07-29	Massachusetts		Andover		YU

1

[Download](#) result set in CSV format.

Export

Export data for import into Symbiota. Only unflagged specimens that have images on iPlant will be exported. After export, records will no longer be accessible via this website.

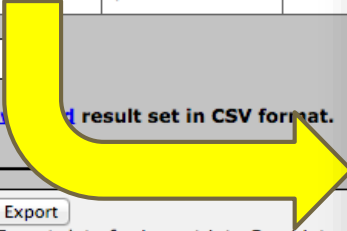
# WORKFLOW: 1° DIGITIZATION - LIGHT BOX - DATA CAPTURE

<a href="#">YU.074948</a>	Epilobium ciliatum Raf.	C. W.	<h3>NEVP Specimen Record Editing Form</h3> <p>You are logged in as: <a href="#">Patrick Sweeney</a>   <a href="#">LOGOUT</a>   <a href="#">SPECIMEN TABLE</a>   <a href="#">COLLECTOR TABLE</a>   <a href="#">NEW RECORD</a>   <a href="#">QUICK FORM</a></p> <p>Don't forget to logout when your session is complete. If you are not the user listed above, please logout.</p> <hr/> <p><b>SPECIMEN ID: 4680   BARCODE: YU.074997</b></p> <p>Barcode: <input type="text" value="YU.074997"/></p> <p>Collection Code: <input type="text" value="YU"/></p> <hr/> <p><b>SCIENTIFIC NAME DATA</b></p> <p>Scientific Name: <input type="text" value="Epilobium palustre L."/>          UUID: c8ce33f7-6157-4004-944a-bcf49cd37828</p> <p>Genus: <input type="text" value="Epilobium"/>          Specific Epithet: <input type="text" value="palustre"/>          Infraspecific Rank: <input type="text"/>          Infraspecific Epithet: <input type="text"/>          Scientific Name Author: <input type="text" value="L."/>          Identification Qualifier: <input type="text"/></p> <p><a href="#">HELP</a></p> <hr/> <p><b>COLLECTOR DATA</b></p> <p>Collector: <input type="text" value="C. W. Swan"/>          CollectorId: 53ceb387-d64e-11e3-992a-005056a70860</p> <p>Collection Number: <input type="text"/>          Verbatim Collection Date: <input type="text"/>          Collection Date (Earliest): <input type="text" value="1882-07-29"/>          Collection Date (Latest): <input type="text" value="0000-00-00"/></p> <p><a href="#">HELP</a></p>		YU
<a href="#">YU.074947</a>	Epilobium ciliatum Raf.	C. W.		W Boston flats	YU
<a href="#">YU.074959</a>	Epilobium coloratum Biehler	Swan,		Graniteville	YU
<a href="#">YU.074967</a>	Epilobium hirsutum L.	C. W.			YU
<a href="#">YU.074985</a>	Epilobium leptophyllum Raf.	C. W.			YU
<a href="#">YU.074997</a>	Epilobium palustre L.	C. W.			YU

1

[Download result set in CSV format.](#)

[Export](#)  
Export data for import into Symbiota.







# WORKFLOW: 1° DIGITIZATION

- **Progress:**
  - 200,000 records




# WORKFLOW: SECONDARY DIGITIZATION

- The secondary digitization stage will involve humans capturing (via keystroking) habitat and phenology data from images of specimens and labels.

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

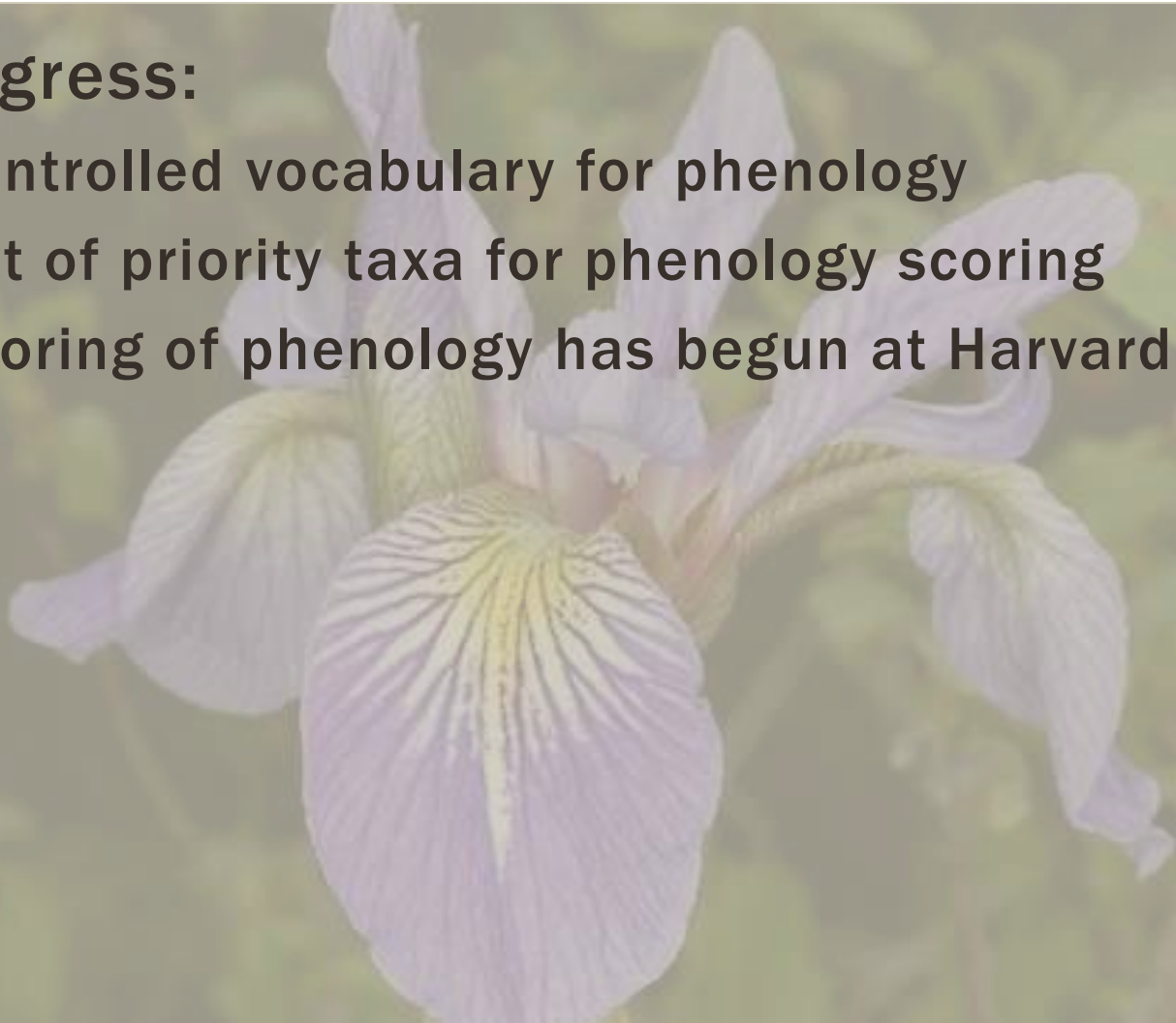
**Label Processing**



# WORKFLOW: SECONDARY DIGITIZATION

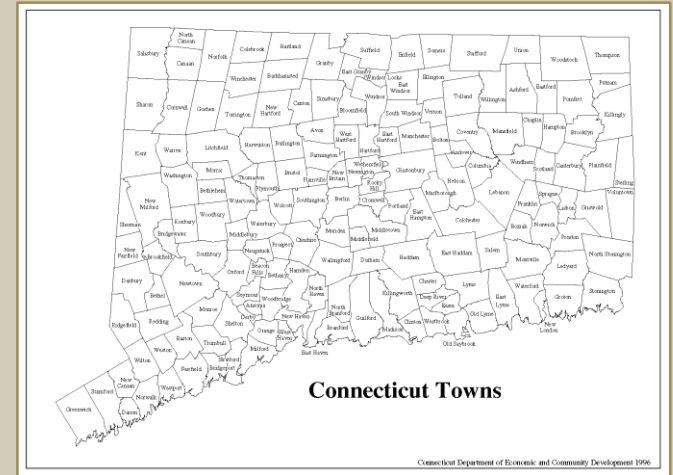
## ■ Progress:

- controlled vocabulary for phenology
- list of priority taxa for phenology scoring
- scoring of phenology has begun at Harvard



# WORKFLOW: SECONDARY DIGITIZATION

- Georeference to at least town level
- Town-level New England gazetteer created
  - obtained GIS town layers from authoritative state source
  - queries in PostGIS to obtain polygon centroids & radius uncertainty



# OUTREACH

## ■ New England Leaf Out Project (NELOP) - Richard Primack (Boston University) & Libby Ellwood



Field Station Concordia and the New England Leaf Out Project have teamed up on a citizen science project to collect leaf out times.

We are investigating the effects of climate change on the tree species of New England. Using both remote sensing and direct observations, we will monitor leaf out times across the region, and whether trees leaf out earlier now than they did in the past due to warming temperatures.

We hope you will help us gather observations of leaf out times this spring to add to the available database of current and historical observations.

If you live in Maine, New Hampshire, Vermont, Massachusetts, Rhode Island or Connecticut, all you need to do is:

- Find one or more trees from the list below in a place that you visit regularly
- This spring, starting in mid April in Connecticut and late April elsewhere, check your tree every couple of days, and look for the first signs of leaf out
- Submit your observations [here](#)
- Please include any information about the tree's environment (high on a mountain, urban street, etc.)

[Enter your observations](#)

### What is leaf out?

**Leaf out:** Please record the date that you first see one, or up to several, new leaves on the tree. In this study, we count a new leaf when it has mostly emerged from the bud and its final shape is mostly visible. This observation should be made in spring when the first leaves are emerging. These young leaves often have a soft or translucent quality to them, and may not yet be green. Please do not report observations of leaves that have reached their full size.



Red Oak, © Richard Primack

# ACKNOWLEDGEMENTS



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



Symbiota Project



FilteredPush

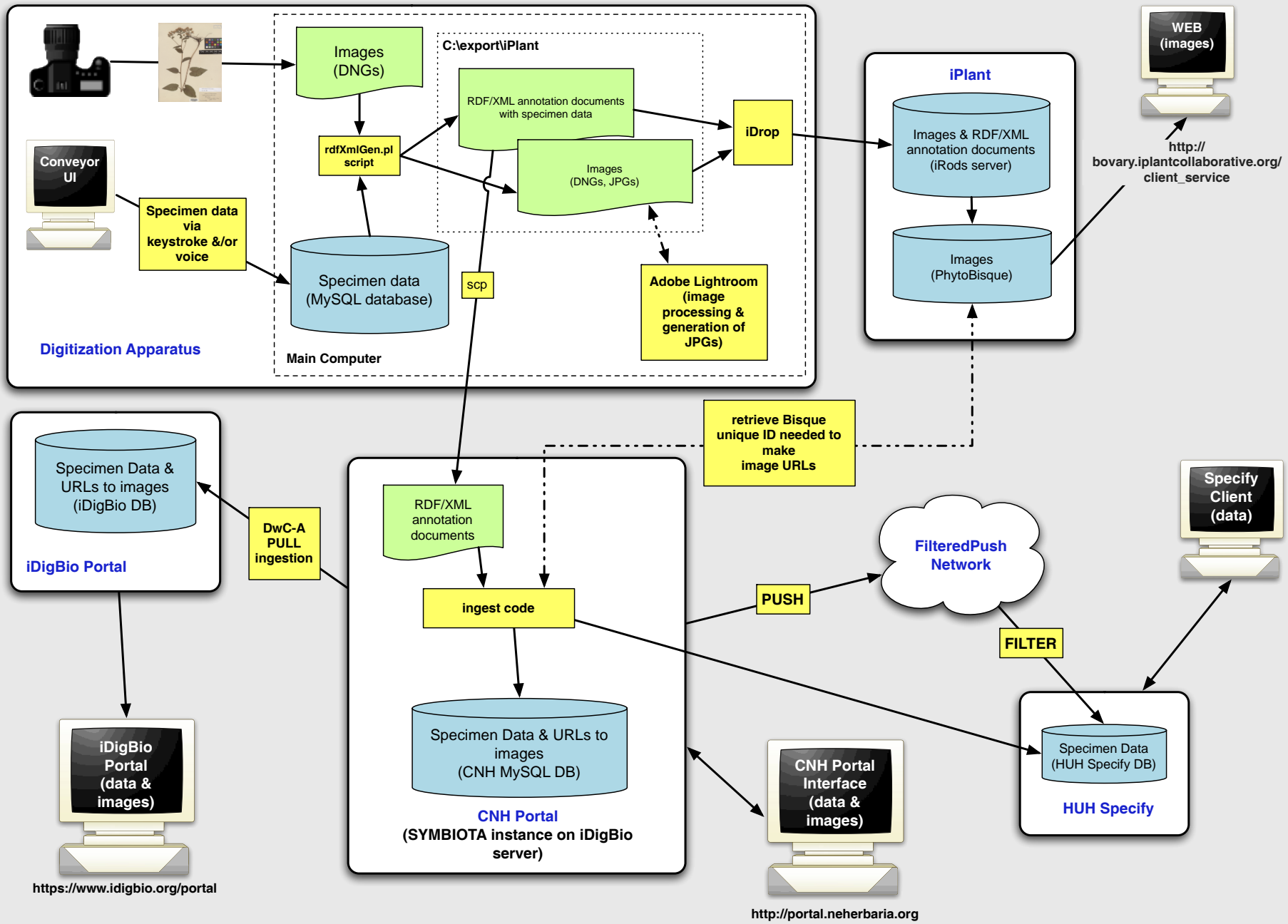


iPlant Collaborative™ *Empowering A New Plant Biology*



Biota of North America

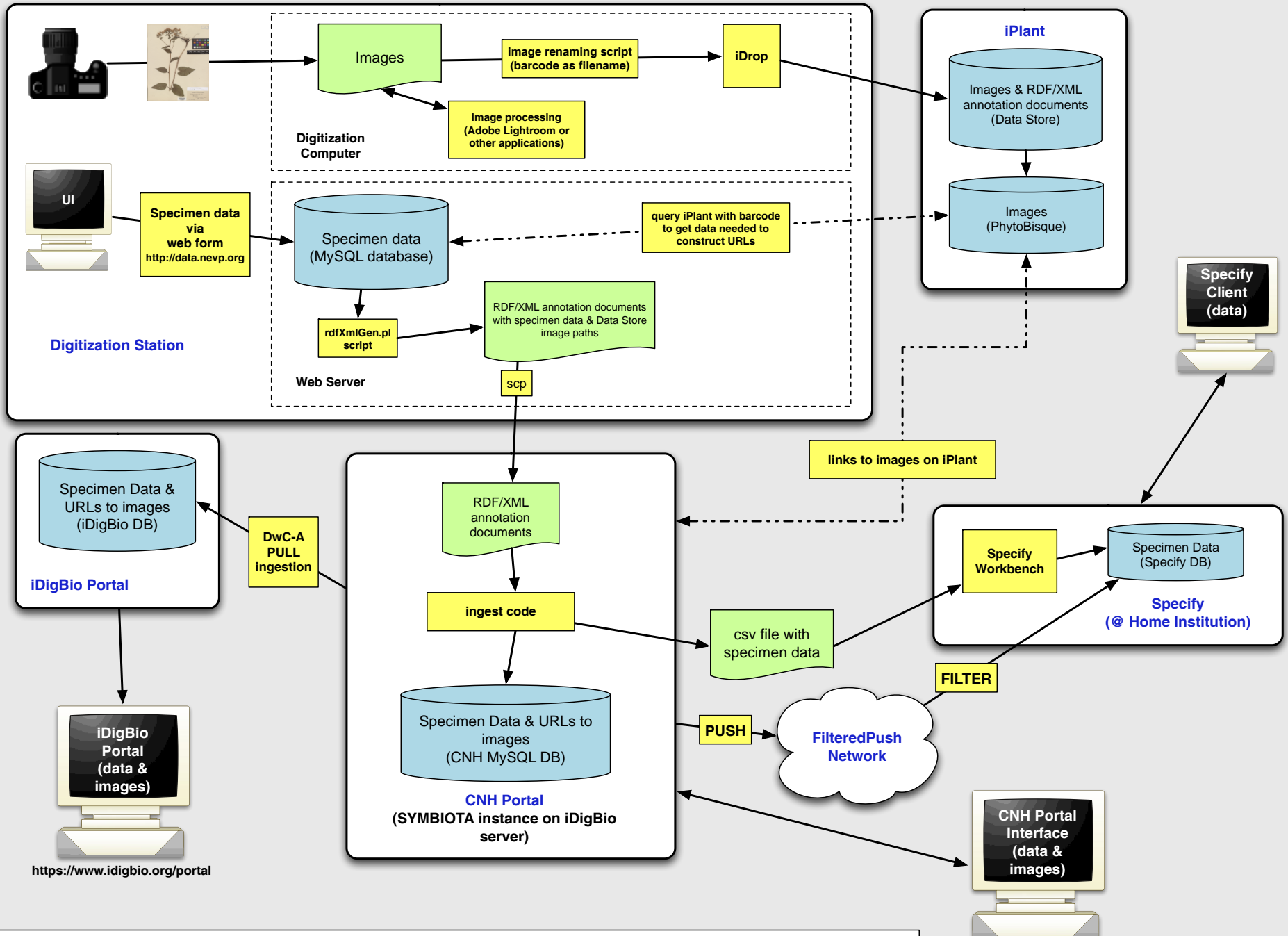
BONAP



NEVP data flow diagram, HUH simplified (2014-03-28, vers. 3)

process (yellow box)   data (green box)   storage (blue box)





NEVP data flow diagram, web GUI, simplified (2014-05-27, vers. 1)

process (yellow box)   data (green box)   storage (blue cylinder)

