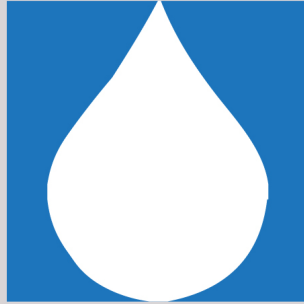


# Mobilizing New England Vascular Plant Specimen Data



to Track Environmental Changes

**P.W. Sweeney, 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**

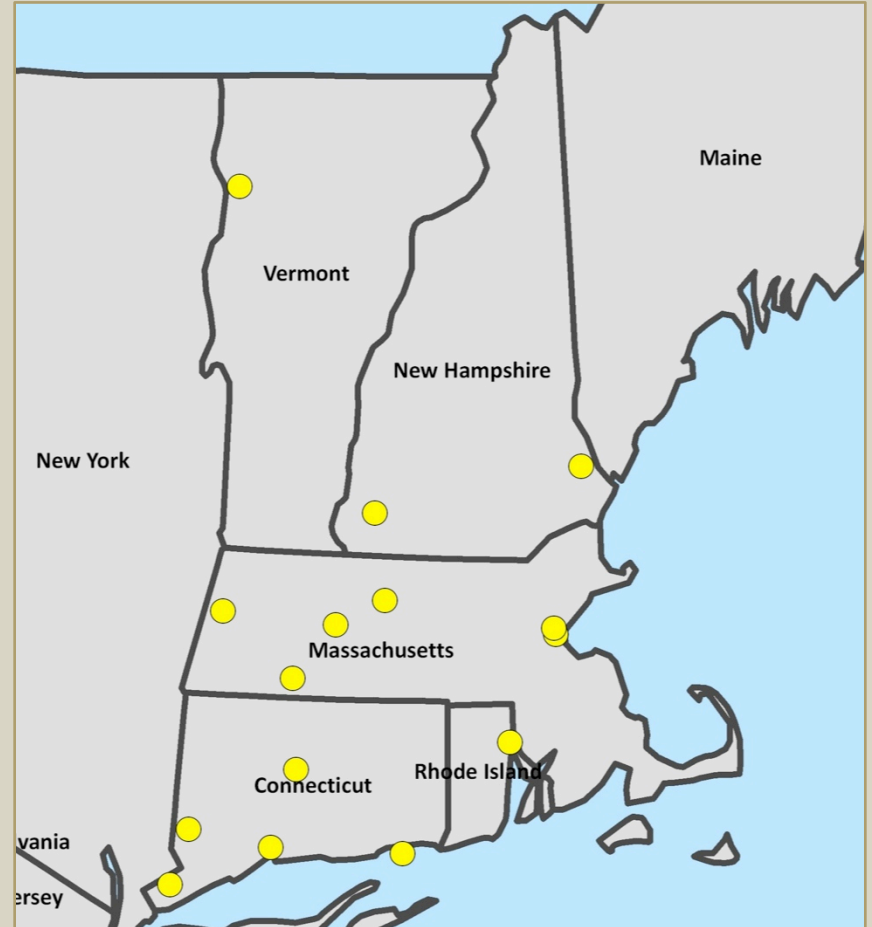


# OUTLINE

- Objectives
- Rationale
- Digitization Plan
  - Organizational Structure
  - Workflow
- Mobilization
- Training & Outreach
- Timeline

# 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 three centuries

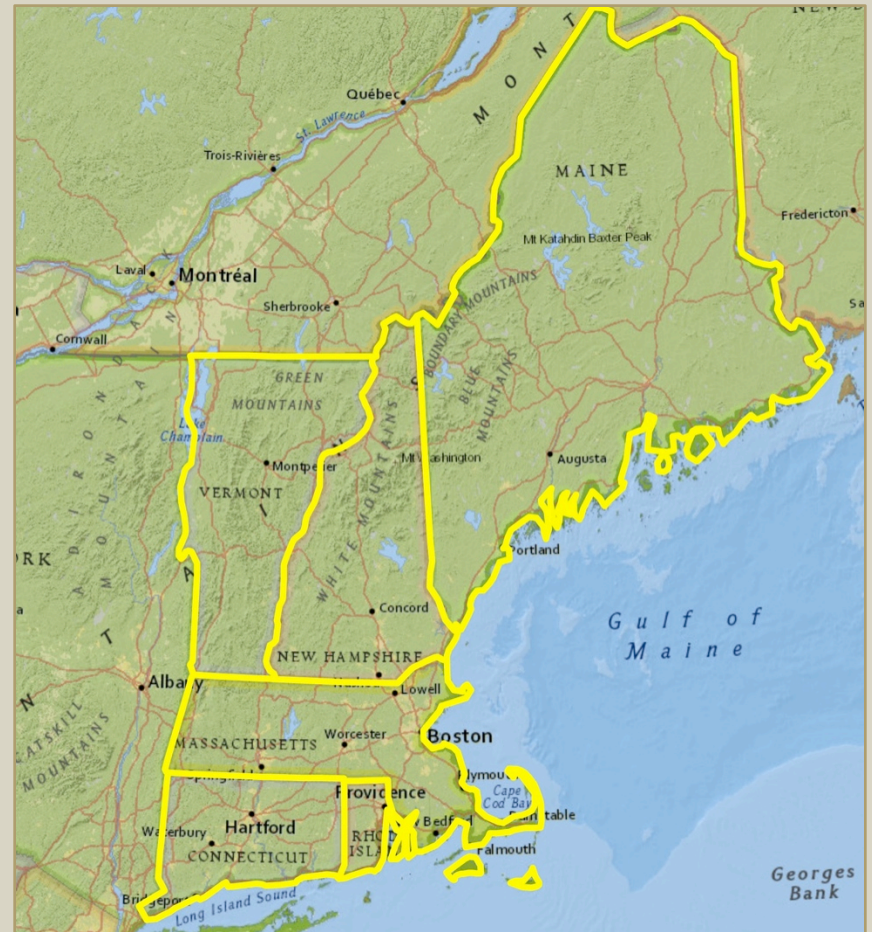




# RATIONALE: NEW ENGLAND

## Why New England?

- Herbaria
- Landscape
- History
- Scientists



# RATIONALE: THEME

## Climate Change

- Plant phenological observations play an important role in the effort to understand the effects of rising temperatures
- Our project will cater to climate-change studies by capturing phenology (flowering and leafing-out stage) data
- We will develop controlled vocabularies for flowering stage and leafing-out stage



# RATIONALE: THEME

## Land-use history

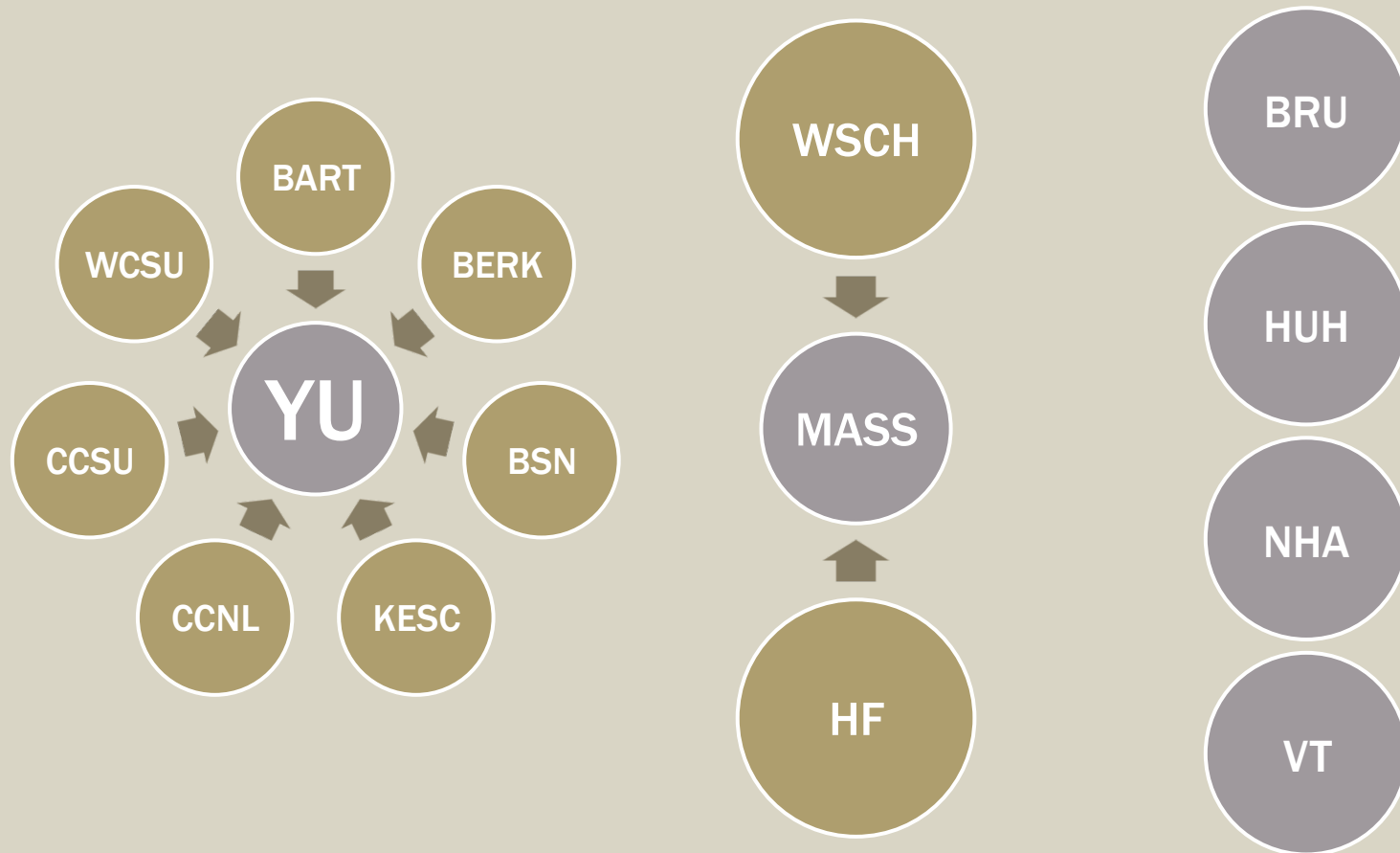
- Herbarium specimens, representing thousands of species with diverse ecological ranges and tolerances can add critical insights into the long-term consequences of past land-use
- We will capture habitat data for a subset of taxa targeted for their particular importance to land-use studies, and we will develop a controlled vocabulary for habitat

# DIGITIZATION PLAN: ORGANIZATION

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



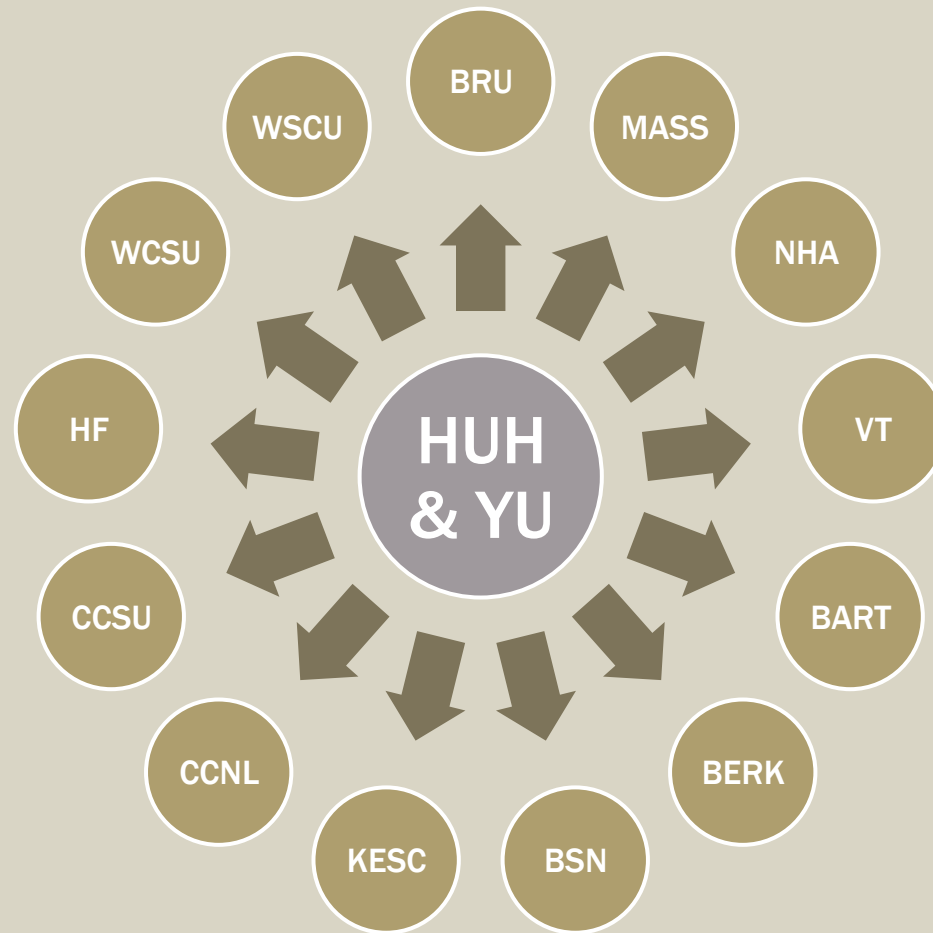
# DIGITIZATION PLAN: ORGANIZATION



# DIGITIZATION PLAN: ORGANIZATION



# DIGITIZATION PLAN: ORGANIZATION



# DIGITIZATION PLAN: WORKFLOW

Collection  
Preparation (Pre-  
capture)



Primary  
digitization



Data  
Enhancement  
(Secondary  
Digitization)

# WORKFLOW: PRE-CAPTURE

- In herbaria, sheets are often organized into folders by genus or species, with a single identification applying to all specimens in that folder. Often organized by geography, as well
- Capture data 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





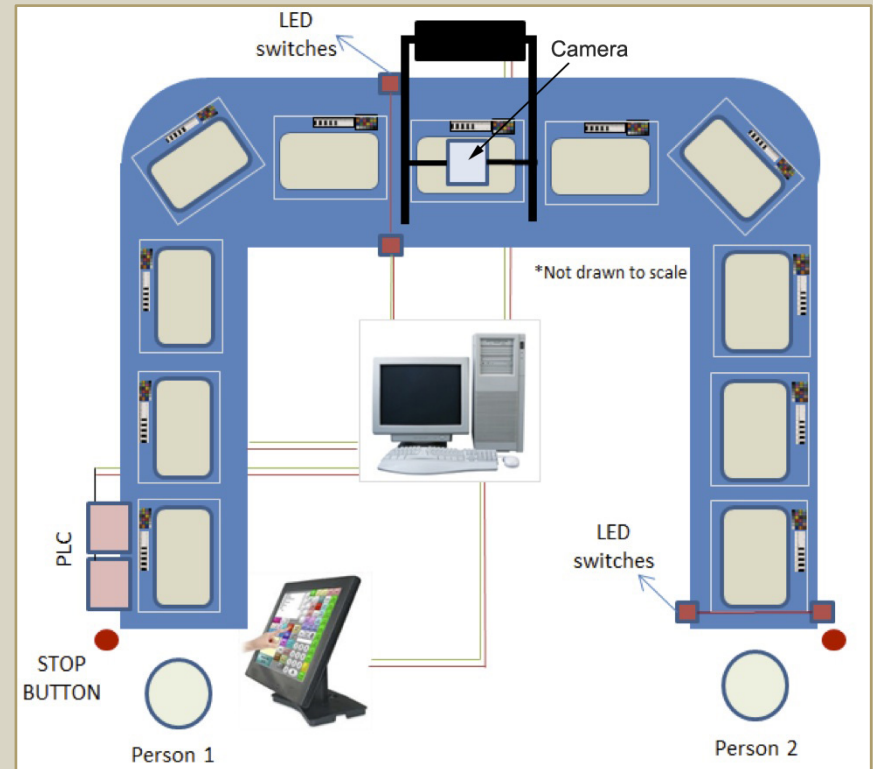
# WORKFLOW: PRIMARY DIGITIZATION

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

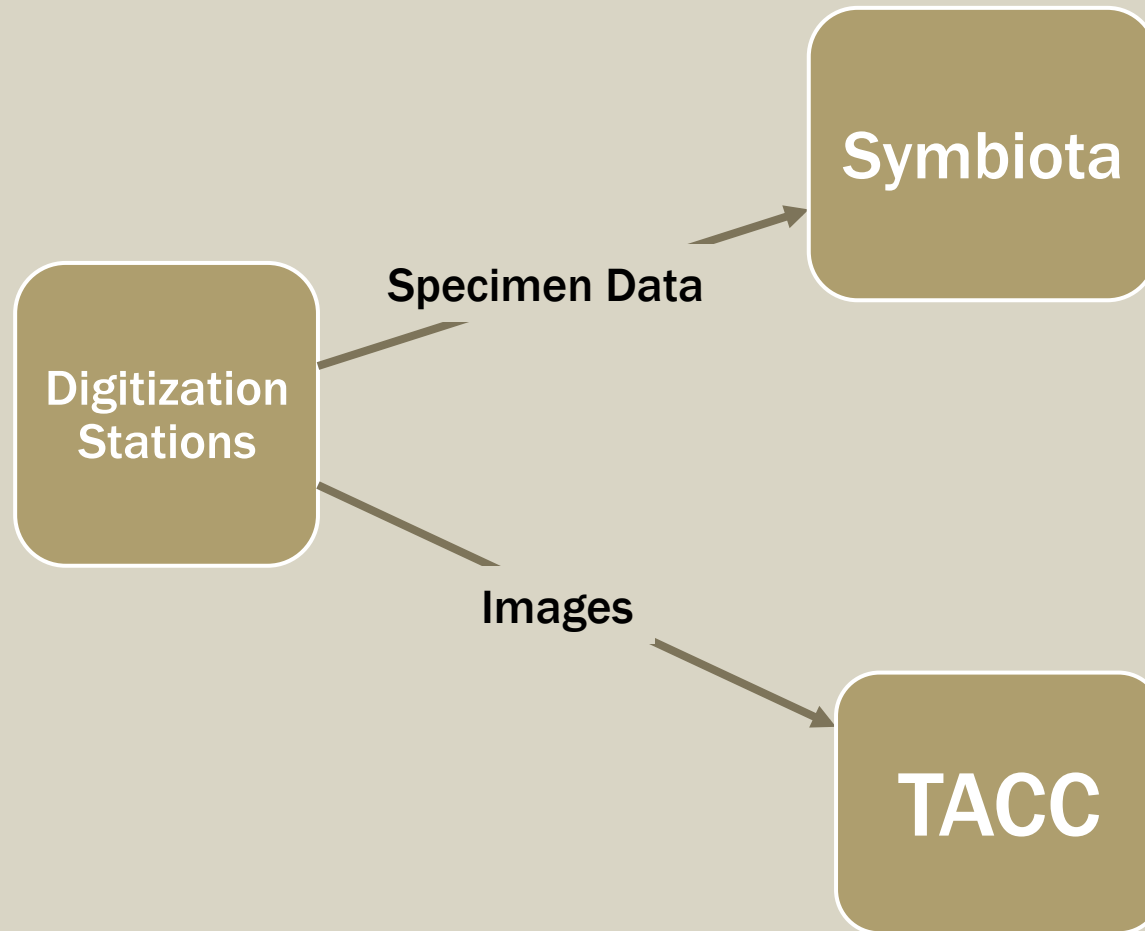


# WORKFLOW: PRIMARY DIGITIZATION

- To increase the efficiency of capturing an image and specimen-level data, we will develop and test an high throughput digitization apparatus



# WORKFLOW: PRIMARY DIGITIZATION



# 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

Collector Info

Collector

Number

Date

Dupes

Occurrence ID ?

Collector

Number

Date

Dupes

Associated Collectors

Other Catalog Numbers ?

Scientific Name

Author

Family

Scientific Name

Author

Family

Qualifier ?

Family

Identified By

Date Identified

Locality

Country

State/Province

County

Municipality

Locality

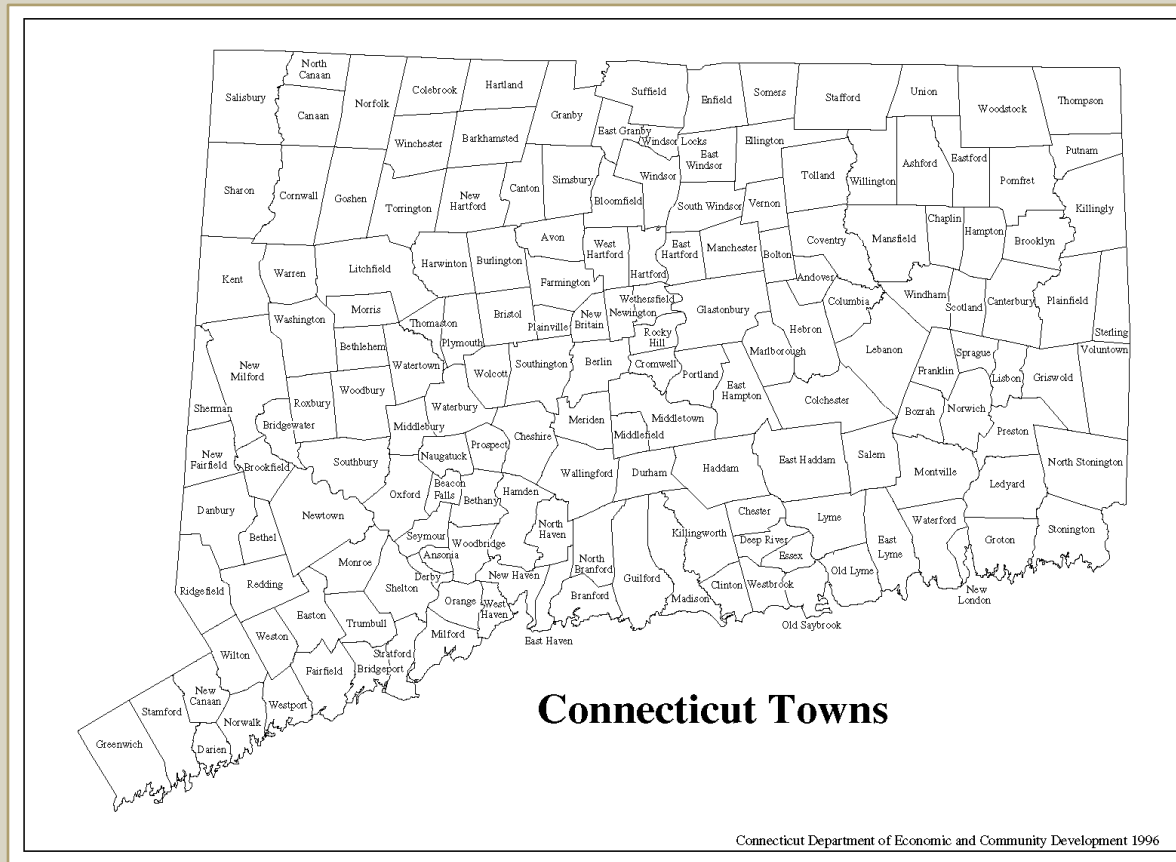
Locality Security

Label Processing



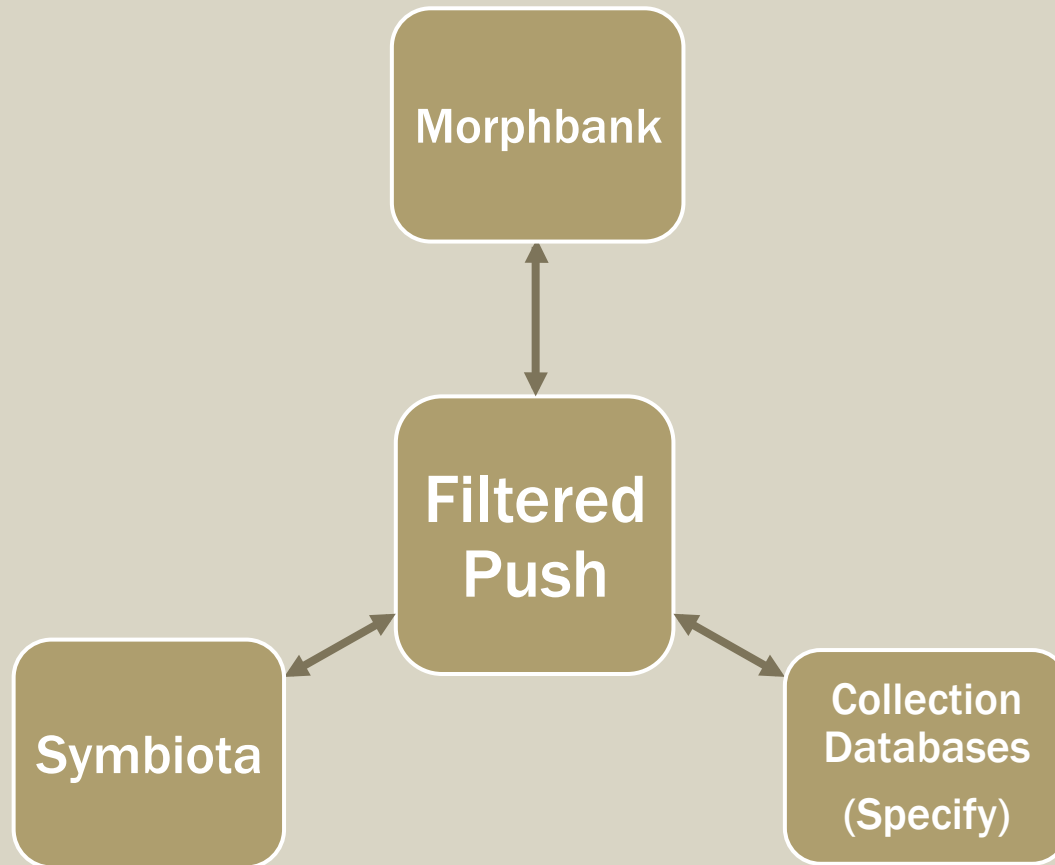
# WORKFLOW: SECONDARY DIGITIZATION

## ■ Georeference to at least town level





# WORKFLOW: ENHANCEMENT



# MOBILIZATION

- All of the data and images will be available to public through *Consortium of Northeastern Herbaria* (CNH) Symbiota portal ([neherbaria.org/CNH](http://neherbaria.org/CNH))

Membership | Governance | Meetings | Resources

## Herbarium Specimen Data Sharing Portal for CNH

Number of records in database: **409883**

### About:

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

Use of any specimen data and related material (e.g., images, species checklists, etc.) on the 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 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 is an NSF funded endeavor based at Arizona State University. We are particularly indebted to ASU for assisting CNH in implementing this Symbiota instance.



AccessionNumber	AccessionNumber	AccessionNumber	AccessionNumber	AccessionNumber	AccessionNumber
CON00000001	CON00000002	CON00000003	CON00000004	CON00000005	CON00000006
CON00000007	CON00000008	CON00000009	CON00000010	CON00000011	CON00000012
CON00000013	CON00000014	CON00000015	CON00000016	CON00000017	CON00000018
CON00000019	CON00000020	CON00000021	CON00000022	CON00000023	CON00000024
CON00000025	CON00000026	CON00000027	CON00000028	CON00000029	CON00000030
CON00000031	CON00000032	CON00000033	CON00000034	CON00000035	CON00000036
CON00000037	CON00000038	CON00000039	CON00000040	CON00000041	CON00000042
CON00000043	CON00000044	CON00000045	CON00000046	CON00000047	CON00000048
CON00000049	CON00000050	CON00000051	CON00000052	CON00000053	CON00000054
CON00000055	CON00000056	CON00000057	CON00000058	CON00000059	CON00000060
CON00000061	CON00000062	CON00000063	CON00000064	CON00000065	CON00000066
CON00000067	CON00000068	CON00000069	CON00000070	CON00000071	CON00000072
CON00000073	CON00000074	CON00000075	CON00000076	CON00000077	CON00000078
CON00000079	CON00000080	CON00000081	CON00000082	CON00000083	CON00000084
CON00000085	CON00000086	CON00000087	CON00000088	CON00000089	CON00000090
CON00000091	CON00000092	CON00000093	CON00000094	CON00000095	CON00000096
CON00000097	CON00000098	CON00000099	CON00000100	CON00000101	CON00000102
CON00000103	CON00000104	CON00000105	CON00000106	CON00000107	CON00000108
CON00000109	CON00000110	CON00000111	CON00000112	CON00000113	CON00000114
CON00000115	CON00000116	CON00000117	CON00000118	CON00000119	CON00000120
CON00000121	CON00000122	CON00000123	CON00000124	CON00000125	CON00000126
CON00000127	CON00000128	CON00000129	CON00000130	CON00000131	CON00000132
CON00000133	CON00000134	CON00000135	CON00000136	CON00000137	CON00000138
CON00000139	CON00000140	CON00000141	CON00000142	CON00000143	CON00000144
CON00000145	CON00000146	CON00000147	CON00000148	CON00000149	CON00000150
CON00000151	CON00000152	CON00000153	CON00000154	CON00000155	CON00000156
CON00000157	CON00000158	CON00000159	CON00000160	CON00000161	CON00000162
CON00000163	CON00000164	CON00000165	CON00000166	CON00000167	CON00000168
CON00000169	CON00000170	CON00000171	CON00000172	CON00000173	CON00000174
CON00000175	CON00000176	CON00000177	CON00000178	CON00000179	CON00000180
CON00000181	CON00000182	CON00000183	CON00000184	CON00000185	CON00000186
CON00000187	CON00000188	CON00000189	CON00000190	CON00000191	CON00000192
CON00000193	CON00000194	CON00000195	CON00000196	CON00000197	CON00000198
CON00000199	CON00000200	CON00000201	CON00000202	CON00000203	CON00000204
CON00000205	CON00000206	CON00000207	CON00000208	CON00000209	CON00000210
CON00000211	CON00000212	CON00000213	CON00000214	CON00000215	CON00000216
CON00000217	CON00000218	CON00000219	CON00000220	CON00000221	CON00000222
CON00000223	CON00000224	CON00000225	CON00000226	CON00000227	CON00000228
CON00000229	CON00000230	CON00000231	CON00000232	CON00000233	CON00000234
CON00000235	CON00000236	CON00000237	CON00000238	CON00000239	CON00000240
CON00000241	CON00000242	CON00000243	CON00000244	CON00000245	CON00000246
CON00000247	CON00000248	CON00000249	CON00000250	CON00000251	CON00000252
CON00000253	CON00000254	CON00000255	CON00000256	CON00000257	CON00000258
CON00000259	CON00000260	CON00000261	CON00000262	CON00000263	CON00000264
CON00000265	CON00000266	CON00000267	CON00000268	CON00000269	CON00000270
CON00000271	CON00000272	CON00000273	CON00000274	CON00000275	CON00000276
CON00000277	CON00000278	CON00000279	CON00000280	CON00000281	CON00000282
CON00000283	CON00000284	CON00000285	CON00000286	CON00000287	CON00000288
CON00000289	CON00000290	CON00000291	CON00000292	CON00000293	CON00000294
CON00000295	CON00000296	CON00000297	CON00000298	CON00000299	CON00000300
CON00000301	CON00000302	CON00000303	CON00000304	CON00000305	CON00000306
CON00000307	CON00000308	CON00000309	CON00000310	CON00000311	CON00000312
CON00000313	CON00000314	CON00000315	CON00000316	CON00000317	CON00000318
CON00000319	CON00000320	CON00000321	CON00000322	CON00000323	CON00000324
CON00000325	CON00000326	CON00000327	CON00000328	CON00000329	CON00000330
CON00000331	CON00000332	CON00000333	CON00000334	CON00000335	CON00000336
CON00000337	CON00000338	CON00000339	CON00000340	CON00000341	CON00000342
CON00000343	CON00000344	CON00000345	CON00000346	CON00000347	CON00000348
CON00000349	CON00000350	CON00000351	CON00000352	CON00000353	CON00000354
CON00000355	CON00000356	CON00000357	CON00000358	CON00000359	CON00000360
CON00000361	CON00000362	CON00000363	CON00000364	CON00000365	CON00000366
CON00000367	CON00000368	CON00000369	CON00000370	CON00000371	CON00000372
CON00000373	CON00000374	CON00000375	CON00000376	CON00000377	CON00000378
CON00000379	CON00000380	CON00000381	CON00000382	CON00000383	CON00000384
CON00000385	CON00000386	CON00000387	CON00000388	CON00000389	CON00000390
CON00000391	CON00000392	CON00000393	CON00000394	CON00000395	CON00000396
CON00000397	CON00000398	CON00000399	CON00000400	CON00000401	CON00000402
CON00000403	CON00000404	CON00000405	CON00000406	CON00000407	CON00000408
CON00000409	CON00000410	CON00000411	CON00000412	CON00000413	CON00000414
CON00000415	CON00000416	CON00000417	CON00000418	CON00000419	CON00000420
CON00000421	CON00000422	CON00000423	CON00000424	CON00000425	CON00000426
CON00000427	CON00000428	CON00000429	CON00000430	CON00000431	CON00000432
CON00000433	CON00000434	CON00000435	CON00000436	CON00000437	CON00000438
CON00000439	CON00000440	CON00000441	CON00000442	CON00000443	CON00000444
CON00000445	CON00000446	CON00000447	CON00000448	CON00000449	CON00000450
CON00000451	CON00000452	CON00000453	CON00000454	CON00000455	CON00000456
CON00000457	CON00000458	CON00000459	CON00000460	CON00000461	CON00000462
CON00000463	CON00000464	CON00000465	CON00000466	CON00000467	CON00000468
CON00000469	CON00000470	CON00000471	CON00000472	CON00000473	CON00000474
CON00000475	CON00000476	CON00000477	CON00000478	CON00000479	CON00000480
CON00000481	CON00000482	CON00000483	CON00000484	CON00000485	CON00000486
CON00000487	CON00000488	CON00000489	CON00000490	CON00000491	CON00000492
CON00000493	CON00000494	CON00000495	CON00000496	CON00000497	CON00000498
CON00000499	CON00000500	CON00000501	CON00000502	CON00000503	CON00000504
CON00000505	CON00000506	CON00000507	CON00000508	CON00000509	CON00000510
CON00000511	CON00000512	CON00000513	CON00000514	CON00000515	CON00000516
CON00000517	CON00000518	CON00000519	CON00000520	CON00000521	CON00000522
CON00000523	CON00000524	CON00000525	CON00000526	CON00000527	CON00000528
CON00000529	CON00000530	CON00000531	CON00000532	CON00000533	CON00000534
CON00000535	CON00000536	CON00000537	CON00000538	CON00000539	CON00000540
CON00000541	CON00000542	CON00000543	CON00000544	CON00000545	CON00000546
CON00000547	CON00000548	CON00000549	CON00000550	CON00000551	CON00000552
CON00000553	CON00000554	CON00000555	CON00000556	CON00000557	CON00000558
CON00000559	CON00000560	CON00000561	CON00000562	CON00000563	CON00000564
CON00000565	CON00000566	CON00000567	CON00000568	CON00000569	CON00000570
CON00000571	CON00000572	CON00000573	CON00000574	CON00000575	CON00000576
CON00000577	CON00000578	CON00000579	CON00000580	CON00000581	CON00000582
CON00000583	CON00000584	CON00000585	CON00000586	CON00000587	CON00000588
CON00000589	CON00000590	CON00000591	CON00000592	CON00000593	CON00000594
CON00000595	CON00000596	CON00000597	CON00000598	CON00000599	CON00000600
CON00000601	CON00000602	CON00000603	CON00000604	CON00000605	CON00000606
CON00000607	CON00000608	CON00000609	CON00000610	CON00000611	CON00000612
CON00000613	CON00000614	CON00000615	CON00000616	CON00000617	CON00000618
CON00000619	CON00000620	CON00000621	CON00000622	CON00000623	CON00000624
CON00000625	CON00000626	CON00000627	CON00000628	CON00000629	CON00000630
CON00000631	CON00000632	CON00000633	CON00000634	CON00000635	CON00000636
CON00000637	CON00000638	CON00000639	CON00000640	CON00000641	CON00000642
CON00000643	CON00000644	CON00000645	CON00000646	CON00000647	CON00000648
CON00000649	CON00000650	CON00000651	CON00000652	CON00000653	CON00000654
CON00000655	CON00000656	CON00000657	CON00000658	CON00000659	CON00000660
CON00000661	CON00000662	CON00000663	CON00000664	CON00000665	CON00000666
CON00000667	CON00000668	CON00000669	CON00000670	CON00000671	CON00000672
CON00000673	CON00000674	CON00000675	CON00000676	CON00000677	CON00000678
CON00000679	CON00000680	CON00000681	CON00000682	CON00000683	CON00000684
CON00000685	CON00000686	CON00000687	CON00000688	CON00000689	CON00000690
CON00000691	CON00000692	CON00000693	CON00000694	CON00000695	CON00000696
CON00000697	CON00000698	CON00000699	CON00000700	CON00000701	CON00000702
CON00000703	CON00000704	CON00000705	CON00000706	CON00000707	CON00000708
CON00000709	CON00000710	CON00000711	CON00000712	CON00000713	CON00000714
CON00000715	CON00000716	CON00000717	CON00000718	CON00000719	CON00000720
CON00000721	CON00000722	CON00000723	CON00000724	CON00000725	CON00000726
CON00000727	CON00000728	CON00000729	CON00000730	CON00000731	CON00000732
CON00000733	CON00000734	CON00000735	CON00000736	CON00000737	CON00000738
CON00000739	CON00000740	CON00000741	CON00000742	CON00000743	CON00000744
CON00000745	CON00000746	CON00000747	CON00000748	CON00000749	CON00000750
CON00000751	CON00000752	CON00000753	CON00000754	CON00000755	CON00000756
CON00000757	CON00000758	CON00000759	CON00000760	CON00000761	CON00000762
CON00000763	CON00000764	CON00000765	CON00000766	CON00000767	CON00000768
CON00000769	CON00000770	CON00000771	CON00000772	CON00000773	CON00000774
CON00000775	CON00000776	CON00000777	CON00000778	CON00000779	CON00000780
CON00000781	CON00000782	CON00000783	CON00000784	CON00000785	CON00000786

# TRAINING & OUTREACH

- Undergraduates and graduate students engaged in digitization & collections activities
- Collection internships to teens



# TRAINING & OUTREACH

- We will establish a network of citizen science observers to provide information on flowering and leafing-out times across the New England region



# TIMELINE

Year 1

Year 2

Year 3

Year 4

Informatics  
Infrastructure

Set-up infrastructure and  
digitization workflows  
Develop & test digitization  
apparatuses at OU

Continued development, testing, and fine- tuning

Pre-capture

Conduct pre- capture stage at primary  
institutions

Pre-capture of small  
collections (at MASS  
& YU)

Primary  
Digitization

Conduct Primary digitization

Secondary  
Digitization

Conduct Secondary digitization.



# ACKNOWLEDGEMENTS

National Science Foundation

iDigBio

Symbiota Project

FilteredPush

Texas Advanced Computing Center (TACC) & The  
iPlant Collaborative