

**Minutes of the Consortium of Northeast Herbarium Meeting  
June 13-14 2014  
The Université de Montréal Biodiversity Centre, Montreal, Quebec**

Minutes taken by Deborah Metsger

**DAY 1**

The 2015 meeting of the Consortium was called to order by Patrick Sweeney at 9:07 AM

**Welcome - Luc Brouillet, Curator, Herbarier Marie-Victorin (MT)**

**1. *CNH: an overview* - Patrick Sweeney, Yale University**

History of Consortium of Northeast Herbaria  
1991, 1992 Association of Northeastern Herbaria  
2004  
2008 - CNH was formed at UMASS  
2009-2014  
Accomplishments include:  
Recruitment of new member institutions  
Symbiota portal  
Annual meetings

Within the North East there are 138 institutions  
24,000,000 specimens  
CNH is comprised of 64 institutions from nine US states and four Canadian provinces

Details on Annual Meeting 2014-06-13:  
There are 20 registrants - 11 from Canada  
15 institutions are represented, covering 8 states and 3 provinces

---

**2. *Accelerating Digitization of Biodiversity Research Specimens through Online Public Participation* - Libby Ellwood, Florida State University**

IDIGBIO

Citizen Science is a way to have people involved and de-myth science

CS = Amateur Science – Civic Science – Crowd sourced Science

Projects in which volunteer's partner with scientist to answer real-world questions

(Cornell)

Open Scientist Definition

Public Participation (History)

- Inadvertent scientist Science, but for a different primary purpose
- e.g. cherry blossom observations Japan 1600<sup>th</sup> and 19<sup>th</sup> centuries Gentleman scientist - self-funded, self-directed science as a hobby e.g. Darwin, Ben Franklin
  
- PPSR Collaborative science between citizen and scientist
- Audubon Christmas Bird Count begun in 1900 - Originally a hunt

**Minutes of the Consortium of Northeast Herbarium Meeting  
June 13-14 2014  
The Université de Montréal Biodiversity Centre, Montreal, Quebec**

Technological Advances – Apps

- New England Leaf Out Project – how has leaf out changed over time  
Herbarium specimens  
1920's spike in collection  
3 billion specimens in biodiversity collections  
There is a need to make them available to researchers tackling contemporary research

IDIGBIO [www.idigbio.org](http://www.idigbio.org)

Digitize

Resources for teachers

Digitizing Biodiversity Specimens

Imaging

Transcribing Specimen Label and Ledger Text

Georeferencing

Annotating

Imaging Blitz - Image 1000's of specimens in a day?

*Digitarium* – on a conveyor belt

**Transcribing specimen labels**

Motivation –

*Atlas of Living Australia*

Quality control – 1 volunteer enter – 1 will verify

*Take Notes from nature* [www.Notesfromnature.org](http://www.Notesfromnature.org)

Presented with only one field at a time – scroll down pick list

4 different volunteers – then find consensus.

Smithsonian

Herbaria@home

Botanical Society of the British Isles

Discover Life - add incentives – more interactive

Optical Character Recognition

Georeferencing

Annotation

Biospex - push the data from image step to other stages – can be pushed through the network by farming

Biospex.org Join our mailing list –

**Minutes of the Consortium of Northeast Herbarium Meeting  
June 13-14 2014  
The Université de Montréal Biodiversity Centre, Montreal, Quebec**

Is there a danger of volunteer pools being depleted? To what extent will they burn out -  
Many people will do one specimen and then stop.

Pros and cons of mass data processing rather than trained technicians –

Level of accuracy decreases substantially

What model would funding agencies prefer?

What kind of training is available?

Badges

Institution based incentives – books at the bookstore

IBC programs

Digitizing Biodiversity Specimens

Biospex

—

BREAK

---

**3. *Symbiota Overview* - Patrick Sweeney, Yale University**

CNH specimen portal

NSF funded endeavour

Crowd sourcing module

Lichen, Bryophytes and Climate Change

Symbiota by default does a big OCR block that can allow you to present data to  
volunteers

Steps

1 – Upload skeletal data

2- Upload images

3- Select records for crowdsourcing

4 Expose records to public

5 Review submissions

<http://neherbaria.org/resources>

Collection Control Panel

- Add new occurrence

Administration Control Panel

**Minutes of the Consortium of Northeast Herbarium Meeting  
June 13-14 2014  
The Université de Montréal Biodiversity Centre, Montreal, Quebec**

Can upload data from a csv file

Guidance on mapping fields available in document

Images

Crowd sourcing control panel

Data protection - Certain data fields are not available for editing by the general user

Exposing to Public

What are the scores?

Download Darwin core archive - search - or download a backup data file. Images can be bundled into the archive.

Geolocate information can be popped right back into Symbiota

User can take it from to be processed to pending review

Duplicate Information from other institutions can be identified, and copied

Direct link to Geolocate

Volunteers create an account

#### **4. *Databasing with Volunteers at MT* - Luc Brouillet, Université de Montréal**

##### **Background**

estimated cost 6\$ / specimen

MT human resources limited

No finance

Students: funding limited (225 hours/yr)

##### **Volunteers**

Friends of the Garden – 30,000

Vast pool of Garden volunteers

Numerous – 150 candidates – 30 integrate

Enthusiastic

Knowledgeable but willing to learn: training

Time commitment usually limited to ½ day per week

Sometimes unpredictable in their commitment

Monitoring required

##### **Nature of Volunteer work and examples**

Divide workflow into small tasks

**Minutes of the Consortium of Northeast Herbarium Meeting  
June 13-14 2014  
The Université de Montréal Biodiversity Centre, Montreal, Quebec**

Allow volunteers to change task

**Tasks**

Herbarium reorganization within families

Obtain an accurate count of the collection

Loans and exchanges

Herbarium specimen digitization

**Projects**

Geo-coordinates of St. Lawrence River islands

Photographing and digitizing

Photographing and minimally digitizing WAT specimens.

Crowd sourcing - Notes from nature

**5. *The Brown University Herbarium: past present, and future*  
Timothy Whitfield, Brown University**

- Stephen Thayer Olney (1812-1878) gentleman scientist
  - Providence businessman
  - Local plants – Catalogue of the plants of Long Island
  - Algae of Rhode Island
  - Leading authority on the genus *Carex*, collected 5-6,000 *Carex* specimens
- August Fenlow
- Charles Wright – Cuba
- Josiah Hale – Southeastern US
- William Whitney Bailey –
  - student of Asa Gray, returned to Brown in 1877
  - incorporated Olney's collection
- James Lawrence Bennett
  - Jeweler in Providence – amateur botanist 1890
- Pringle Mexican,
- McCoun Canada
- James Franklin Collections

TODAY!

New facilities

Incorporation of new specimens from local ecological studies

Specimens from Papua New Guinea

**Minutes of the Consortium of Northeast Herbarium Meeting  
June 13-14 2014  
The Université de Montréal Biodiversity Centre, Montreal, Quebec**

---

**6. *Macroalgal Consortium* Janet Sullivan University of New Hampshire**

49 institutions in 29 states

Georeferencing all specimens

3 different tiers of participation:

- UNH PDC Primary digitizing Centres
- Contributing institutions
- Digitization centres – digitizing own collections.

Student, faculty, administrators

Digitization equipment distributed amongst 18 institutions

50-100 years old 50% older than 50, 10% older than 100 yrs old

Pre-digitization curation

What do we have in our collection?

Barcode placement: Importance of barcode label placement at the bottom of specimen in order to facilitate crowd sourcing and database from labels

Skeletal label capture

Images stored on IDigBio cloud

Macroalgal Herbarium portal <http://macroalgal.org>

Students and others can have editorial privileges

Taxonomic information is imported automatically from Algaebase

Specimen label information is captured

Research Applications

Systematics

Floristic Diversity and Biogeography

Historical Comparisons

Human impact on coastal environments

Invasive species tracking

---

**7. *New England Vascular Plant Digitization Project* - Patrick Sweeney, Yale University**

TCN is to provide data support for the changes in vegetation over the last hundred years

Climate Change

Plant phenological observations

Land use history

**Minutes of the Consortium of Northeast Herbarium Meeting  
June 13-14 2014  
The Université de Montréal Biodiversity Centre, Montreal, Quebec**

Habitat data for a subset of taxa

Partners

Collection Prep – Primary Digitization – Data Enhancement 2ary digitization

Web form creates slip that links to geography

Throughput digitization apparatus

800 specimens a day

Folder levels – Collector, Coll no., Coll date, State, Town.

Light Box – mid-sized collections

Digitization Stations – specimen Data

Images – iPlant

Light Box

Primary digitization progress since Nov. 2013 = ca 100,000 records

Secondary digitization will involve humans capturing digitization records

Georeferencing to at least town level. Town-level New England gazetteer produced.

Digitizing to town centroids

Training Activities

New England Leaf Out Project

Chuck Davis - Curio

Score phenology on herbarium specimens

YUBio Data Portal - on campus recording plant occurrence

---

**8. *Filtered Push*, James Macklin , Agriculture and Agrifood Canada**

FP-Data Entry

Duplicate Specimens – a classic

Collecting events zoology

Challenge – Improve data capture efficiency without sacrificing quality

Requirements

Duplicate/related query result must be presented in near real time

Form must be populated faster than can be typed

Integration

Enter query fields

Look up matches in index

Rapid data entry

**Minutes of the Consortium of Northeast Herbarium Meeting  
June 13-14 2014  
The Université de Montréal Biodiversity Centre, Montreal, Quebec**

**Business Meeting**

Convened at 3:20 PM

Governance

**Current Steering Committee**

Dorothy Allard – University of Vermont

Bob Capers – University of Connecticut

Mike Donoghue –Yale University

Jennifer Doubt – Canadian Museum of Nature

Alina Freire-Fierro – Philadelphia Academy of Natural History

Paul Harwood – Brooklyn Botanic Garden

James Macklin – Agriculture and Agri-Food Canada

Deborah Metsger – Royal Ontario Museum

Chris Neefus University of New Hampshire

Janet Sullivan University of New Hampshire

Paul Harwood's position is vacant because Brooklyn Botanical Garden Herbarium has been closed. Status of Brooklyn Botanical Garden – Susan Peel only person left – NYBG has their data – specimens have been boxed up. President said that the situation is temporary. The crack in the building is too expensive to repair. Paul has moved to the Catskills, Kerry Barringer is Kerry working as an independent contractor.

**2015 meeting** Barbara Thiers – New York

Other topics

- **Crowdsourcing**

Patrick can try to facilitate customization

- **Endangered, Threatened and Special Concern Taxa in portal**

Symbiota allows for masking of data –e.g. all locality information can be masked

Some have proposed national list of rare things?? Not a good method.

Filtering – needs to be done on state level

Locality security on each record in Symbiota

**Discussion on Endangered, Threatened and Special Concern Taxa in portal**

Dorothy – Use Naturereserve listing ranks – rather than state ranks. Anything that is S1 or S2 should be masked.

Mary – GBIF report on how to deal with sensitive species. – 8 years out of date.



**Minutes of the Consortium of Northeast Herbarium Meeting**  
**June 13-14 2014**  
**The Université de Montréal Biodiversity Centre, Montreal, Quebec**

- Patrick – Administrators can see all data. It hasn't been a problem.
- Luc – People should be sent to regional plant data center – they deal with the information legally
- James – Idea of policing is impossible. Data has been published and therefore is already in the public domain. Open data, open science, don't hide anything.
- Dorothy – Should give out data selectively to people that we know will use the data properly – cannot give to everyone.
- Ellen – Need to have good relationship with people in Natural Heritage information Centers. Don't need to kill access – but do need to make it harder for people to access it.
- Marcia – Date and collector information if placed in dups – will be able to be accessed.  
Can fuzz data so that you give location + or – a certain distance.  
Fuzzing – taking coordinates and broaden them  
Redacting – hiding data at a certain level
- Luc – Other side of coin – if people don't know about a rare plant they can't protect them.
- James – Google provides information in the literature. USDA plants provides information but it is ancient.
- Dorothy – Nature serve updates information regularly every few years
- Luc – Specimens published on more than one server – Canadensys + CNEH. It is a political rather than a rational decision.
- Marcia – Has anyone done a study to see if it is really a problem? If flagged as redacted then it might be ok.
- Yukon – Institutions can redact information before it comes OR – portal would run a query that would set a security risk for certain fields.
- Dorothy – Could CNEH design a study based on hits on redacted data?  
Submit a form – you have permission to use this form.
- James – Two Darwin core fields that would have to come out of individual data sets.  
Data Generalization  
Information withheld
- Patrick – Locality security reason Symbiota may build it in within the next month. Need a policy and procedure for which fields would be included - Could you just do a centroid according to town? Thousands of meters uncertainty.

**Minutes of the Consortium of Northeast Herbarium Meeting  
June 13-14 2014  
The Université de Montréal Biodiversity Centre, Montreal, Quebec**

What is NEDP doing? For rare species the image would not be served.

Agree – Use nature serve lists to redact information state by state until further notice.

If someone needs that data they can request the data. IF the request is reasonable the data will be sent to them. Mechanism will be worked out. Patrick is to circulate questionnaire to steering committee.

Dave Boufford – vandals and thieves are only interested in a few plants.

**Other Business**

None

**Meeting adjourned to a tour of MT and the facilities of the Biodiversity Centre**

**Minutes of the Consortium of Northeast Herbarium Meeting  
June 13-14 2014  
The Université de Montréal Biodiversity Centre, Montreal, Quebec**

**Day 2**

***What Is In a Name?* Luc Brouillet, Université de Montréal**

Types of species/ infraspecies names found in collections

Slide 1

Not problematic

Accepted names

Homotypic synonyms of accepted names

Heterotypic synonyms of accepted names acknowledged by specialists

Homo

Slide 2

Problematic

Homonyms in different kingdoms

Controversial heterotypic synonyms

Missed homonyms

In part names

Sensu names

Orthographic variations not listed in database

Published names not listed in databases

Unpublished names

Names with typos ( including total fabrications) ( usually corrected at entry

Names with proper authority but incorrect rank ( var. instead of f., etc.

Anamorphs not associated with teleomorphs ( Fungi)

Authority issues

Cannot assume that databases are correct

Problem – researcher annotates specimen before they complete publication and  
e.g. annotates it as ssp and then publishes as a species.

Binomial with several different authorities only a fraction of which are real.

List typos in notes but correct as it is entered.

**Minutes of the Consortium of Northeast Herbarium Meeting  
June 13-14 2014  
The Université de Montréal Biodiversity Centre, Montreal, Quebec**

**Need to do specimen validation before data entry**

**Name Validation**

Database is main means of validating names

Compliment BHL

Nomenclature databases: IPNI

Names +citations

Without status evaluation

Without synonymy exc. Basionyms + homotypic

Taxonomic databases (often regional):

Names (+citations)

Synonymy

(sources)

Accepted names in Tropicos –

**NAME DATABASE ISSUES**

Extensive: best sources available

Complete – no

Accurate: not always

Contradictory between databases: often

Internally contradictory: sometimes

Usage requires judgement and cannot be fully automated

**Taxonomic Databases Issues**

- Two types
  - specific taxonomic focus ( ex.: ferns)
  - regional focus ( IT IS EuroMap)
- up-to-date: often +/- dated
- sources: not always provided
- congruence between db : not always
- taxonomic traditions/ usages

**Minutes of the Consortium of Northeast Herbarium Meeting  
June 13-14 2014  
The Université de Montréal Biodiversity Centre, Montreal, Quebec**

- taxon concepts

**Responsibility?**

- Will end-users use data if names perceived as having no coherence?
- -no
- Who has the taxonomic expertise?

Who is managing data?

- Collection data managers

But who has tools/resources?

- Integrators
- IT specialists

**Needs of Data managers**

- In collection db, most names are probably not problematic
- IT tools to rapidly identify names that are problematic
- More collaboration to improve names in databases
- International consensus on taxonomy of taxa based on systematic data

**Minutes of the Consortium of Northeast Herbarium Meeting  
June 13-14 2014  
The Université de Montréal Biodiversity Centre, Montreal, Quebec**

***What's in a Name?* James Macklin Agriculture and Agri-Food Canada**

Global Names Architecture

Plant-based resources

Exploring Taxon Concepts

Kurator

GNI Global Names Index

Place to see if name exists at all

GNUB Global names usage bank

Name usages anchored to documentation sources (e.g. published literature)

Global Names recognition and discovery tools and service

INBIO –

Taxon finder

**Global Names Index 18,000,000 names – way to access dirty bucket**

Can see where the incorrect names are coming from

Scientific Names Parser –

Nomenclator

IPNI nomenclature specialists

Can get two or three answers that come from different sources – don't always agree.

Global Names resolution tools and services

Catalogue of Life vs. It is. Vs. ....

Poor spelling:

Tropicos Names and relationships - isn't best at sourcing where the decision is from.

(Global)

Canadensys (Regional)

The Plant List - relies on multiple sources - gives a person's initials

Taxonomic Name Resolution Service (TMNS)

**Minutes of the Consortium of Northeast Herbarium Meeting  
June 13-14 2014  
The Université de Montréal Biodiversity Centre, Montreal, Quebec**

Looks across several sources and gives you a name and scores a confidence level

World Plants - - spits out distribution information – but no source

Classification but not source

ITIS – checklist

ETC: Exploring Taxon Concepts through Analysing Fine-Grained Semantic Mark-up of Taxonomic Literature

**Concept Comparison Challenges**

Concepts published in: journals, monographs, floras/faunas, checklists

Geographically contained:

Vary in amount of detail

Names are not enough

A = A +B

Biologically driven

Qualitative to Quantitative

Access to f

Kepler Kurator

No one-stop shop for names

Difference of opinion between resources (trust/authoritative)

Concept evaluation/visualization in its infancy

All resources do not have web services available, and some not reliable.

---

***Have We Got the Names “ Right”* David Shorthouse, Université de Montreal**

Why, Geekery and Logistics

Canadensys Explorer

1,899,979 records

No attempt to achieve taxonomic consensus

Hemihomonyms same name across codes

Darwin Core Archive Validation

Darwin Core terms

**Minutes of the Consortium of Northeast Herbarium Meeting  
June 13-14 2014  
The Université de Montréal Biodiversity Centre, Montreal, Quebec**

Occurrence Records

Identity, Nomenclature, Taxonomy, Taxon Concepts

Identification

Scientific name linked to an identification.

Scientific name ID - an identifier for the nomenclatural ( not taxonomic) details

Identification Qualifier cf. , aff.

Nomenclature

Name Published name

Original name ( basionym)

Darwin Core extensions

Identification History

Where is there more help

TDWG Listserv

Taxacom Listserv

Canadensys Google Group

Herbaria listserve

Applecore should eventually go to DWC wiki list

So, where, how do we start?

iPhylo

Google refine – open refine

<http://www.canadensys.net/refine/>

use some dynamic searches