Experiences from The New York Botanical Garden and Using KE Software's Electronic Museum (EMu)

History of Specimen Databasing at NY

- 1990 - Plans are made to begin specimen databasing
- 1993 - Work is started on the first specimen database

History of Specimen Databasing at NY

- 1995 - Collaboration with Harvard University on a type specimen database
- 1996 - Data first made available on the Internet

History of Specimen Databasing at NY

- 1998 - Herbarium Imaging Laboratory established
- 2003 - Imaging of approximately 90,000 plant type specimens is completed

History of Specimen Databasing at NY

- 2004 - Data is transferred to KE Software's Electronic Museum
  - New web interface with EMu backend goes live

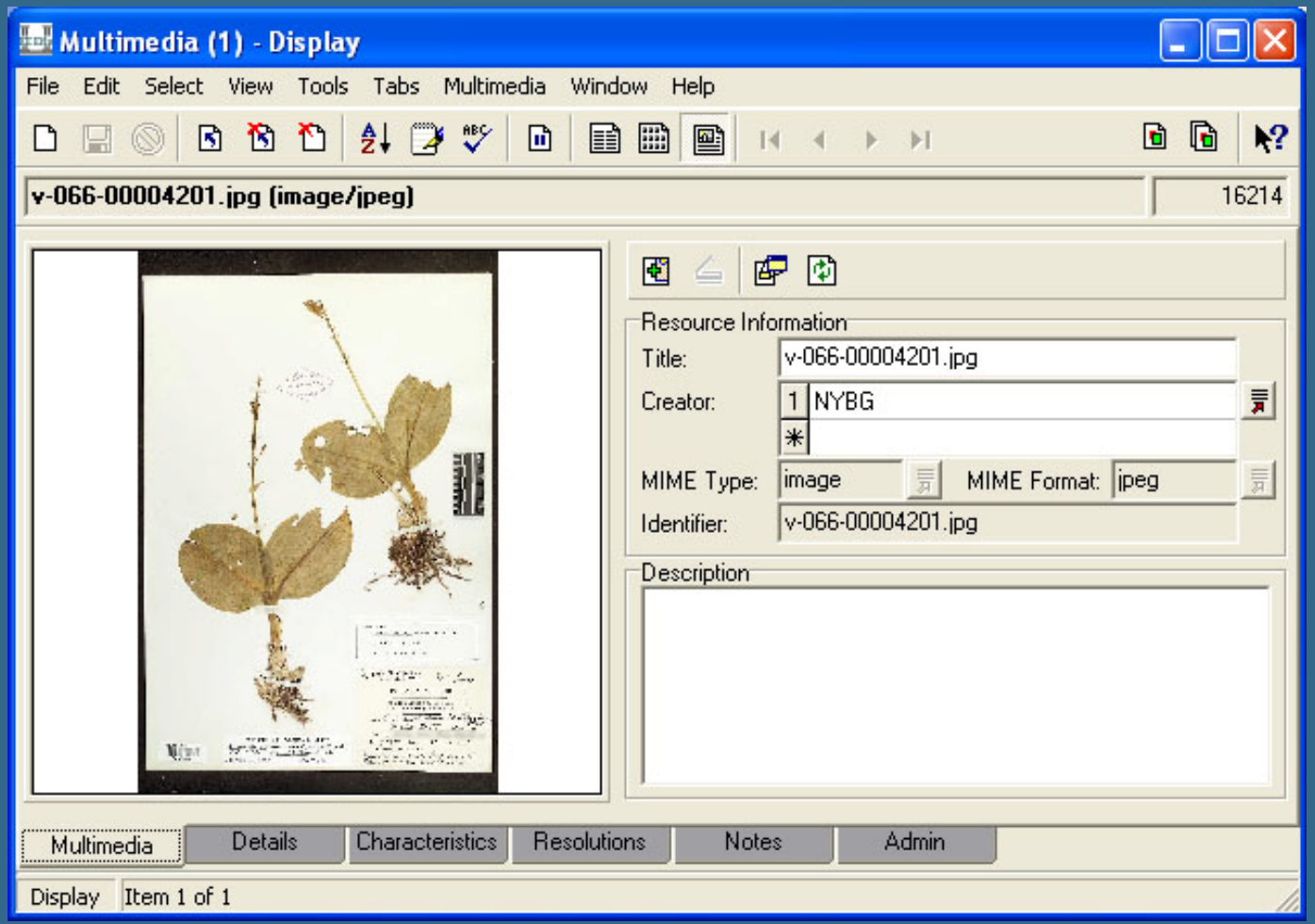
Goals of Herbarium Specimen Databasing

- to make specimen data available electronically for use in biodiversity research projects by NYBG staff and scientists around the world
- to reduce handling of specimens by supplying data transcription and images for uses that do not require direct examination of specimens
- to reunite data elements (e.g., photographs and drawings, manuscripts, published works, microscopic preparations, gene sequences) derived from a specimen with the catalog record for that specimen

EMu Overview

- Digital File Management
  - Any digital file can be attached to any module in the system.
  - Examples include images, documents, mpeg movie files, spreadsheets, etc. (see some examples here).

The Electronic Museum (EMu)



EMu Overview

- Pulls together disparate data elements residing in many physical locations
  - Individual scientists' specimen databases.
  - Institutional databases (frozen DNA samples, economic botany data, pharmaceutical specimen extract data, etc.).
  - Bibliographic data, including published works as well as manuscripts, letters and other unpublished artifacts.

The Electronic Museum (EMu)

- Modular design allows common data to be housed together.
- Modules include Taxonomy, People/Organizations, Bibliography, Loans, Multimedia, and others.
- These modules link to one another so that, for instance, every module has a connection to Multimedia.

Databasing at NY

- Most specimen data entry is done on a project-based basis.
- Specimen catalogers typically have a bachelor's degree in biology and experience with natural history collections.
- Feedback from these catalogers is encouraged and, indeed, crucial to the success of databasing.

Database Management

- Having a database manager to oversee all aspects of the data and the database is critical.
- As a database grows in size and functionality (georeferencing, imaging, web presentation, etc.) more managers of the data have been needed.

Database Management

- Preparation of documents that standardize data entry, a procedural or best-practices document, will greatly reduce the amount of data cleanup that is necessary.
- Non-standardization of data entry can serve as an enormous strain on resources later on.
- Standardization is impressed on catalogers during a training phase.

Data Sharing

- Specimen data downloads from NYBG website.
- DiGIR-formatted data are exported from EMu.
- KE Software plans a set of DiGIR fields in an upcoming version.

Internet Access of EMu

- Specimen data now available online: http://sciweb.nybg.org/science2/VirtualHerbarium.asp
  - Features include checklists, downloads of data, and soon, maps.

Bottlenecks of the Databasing Endeavor

Bottlenecks

- Type Specimens
  - Bibliography
  - Synonymy

These can be considered “interruptions” to databasing, as users need to consult library collections (or, increasingly, online databases).

Bottlenecks

- Specimen Curation
  - Time is needed to correct curating errors discovered during databasing.
    - Misfiled specimens
    - Misidentified specimens
  - Databasing in this context can be thought of as another method of increasing the accuracy of the curation of a collection.

Bottlenecks

- Fieldbook vs. Retroactive databasing
  - Fieldbook data generally much more organized and amenable to computerized data entry.
  - In some cases for specimens already filed, it may be advantageous to arrange them for more efficient data entry.

Bottlenecks

- Transcription of field data to the database adds one more step in making data available.
- Use a portable version of the database in the field.
- Associate images and other supplemental data immediately.

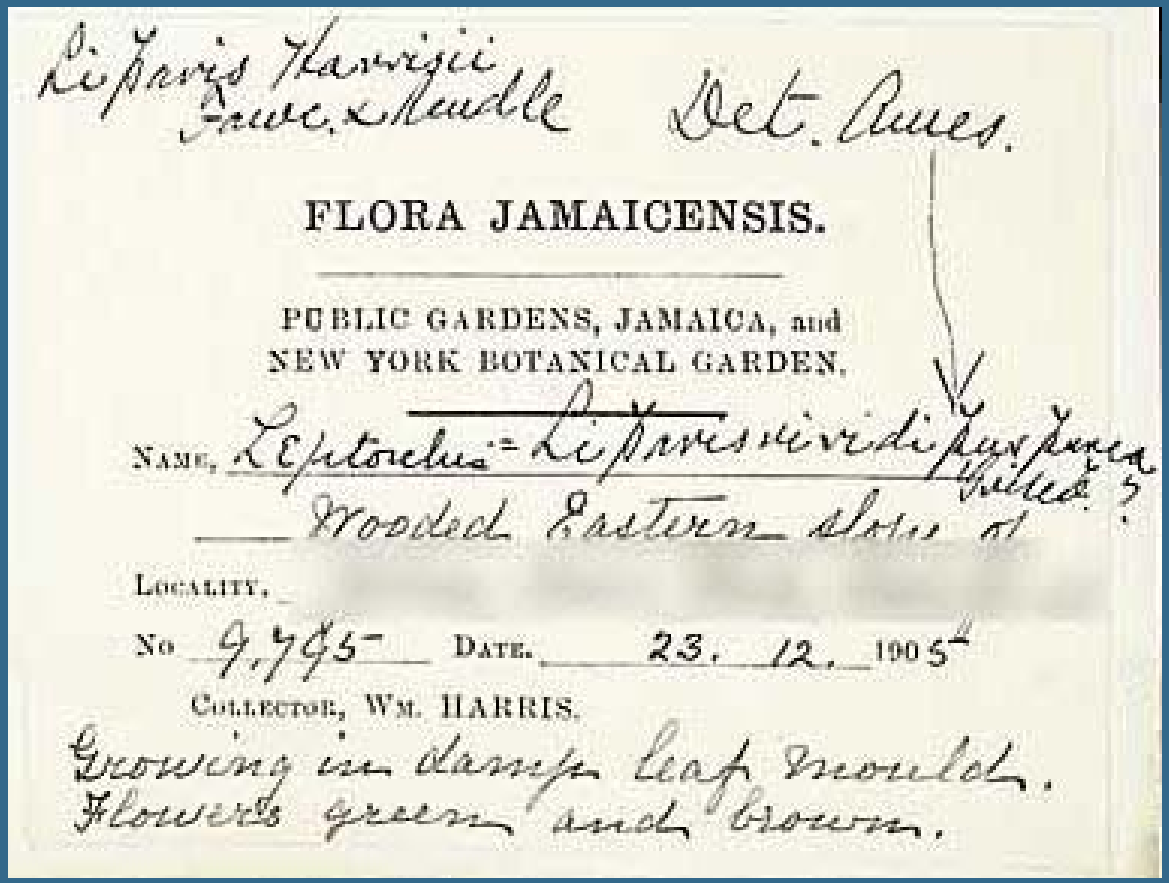
Bottlenecks

- Specimen Georeferencing
  - Older specimens from some parts of the world (in our experience, for example, Latin America) are not well suited for quick georeferencing.

Bottlenecks

- Endangered Species
  - Specimens need to be identified as such.
  - Online access to locality data needs to be restricted.
  - Locality data on any images or documents need to be removed.
  - Generalization of any maps based on locality (specifically, georeferenced specimens).

Endangered Species



Liparis Hawaiii
Swe. x Rendle Det. Ames.

FLORA JAMAICENSIS.

PUBLIC GARDENS, JAMAICA, and
NEW YORK BOTANICAL GARDEN.

NAME, Leptochus = Liparis viridiflora Jamaica
Wooded Eastern slopes of

LOCALITY,

No 9,745 DATE. 23. 12. 1905

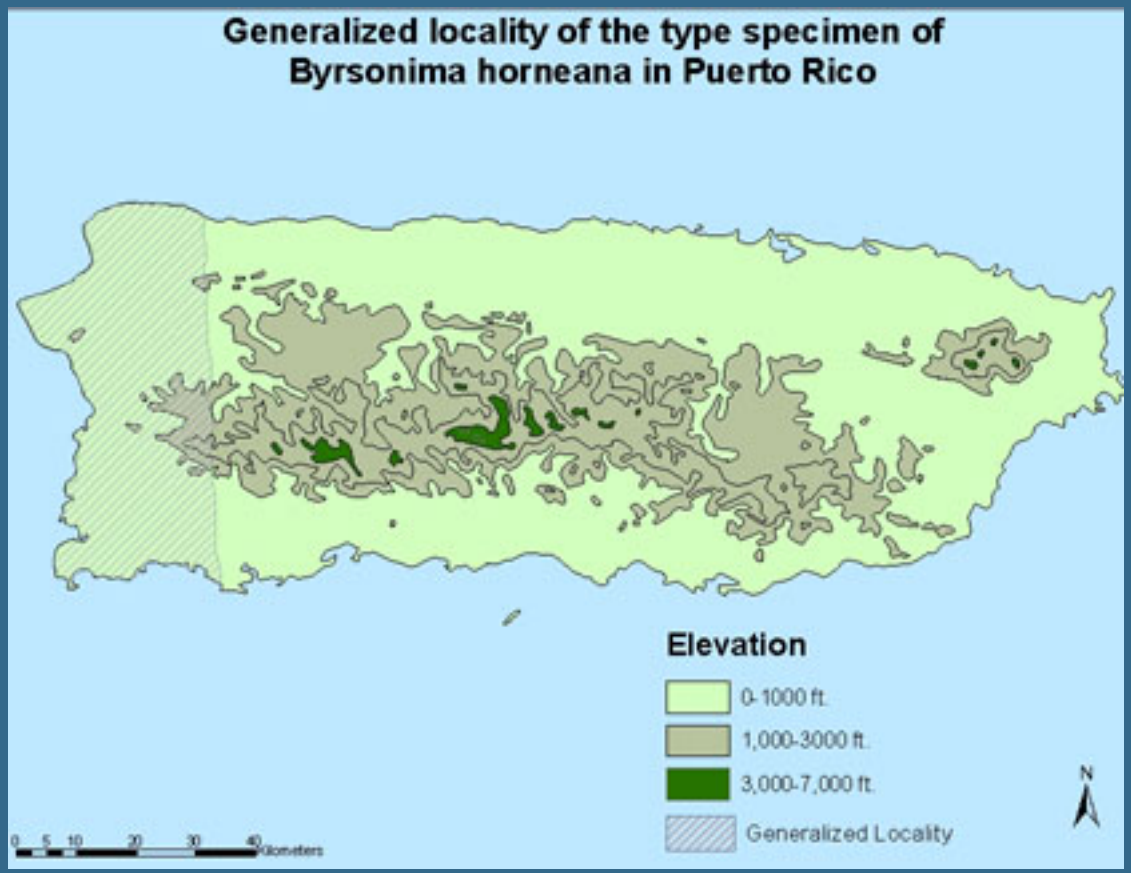
COLLECTOR, WM. HARRIS.

Growing on damp leaf mould.
Flowers green and brown.

Endangered Species

Filed As	**Orchidaceae** Goodyera formosana Rolfe **(isotype)**
All Determinations	Goodyera formosana Rolfe
Location	**China**. Taiwan. **ENDANGERED** This specimen represents an endangered or threatened species. The specific locality has been removed from the on-line record to protect this species from over-collection. These data may be supplied to researchers on request.
Collector	A. Henry 409,
Description	Flower whitish. Flower.
Other	**NY Specimen ID:** 8856 **Collection Notes:** Collection date: 04 Apr

Endangered Species



Steps Toward Minimizing Bottlenecks

- Database Management
- Create procedures and ensure that all users are familiar with them
- Database from a fieldbook if at all possible, and try to capture as much data in the field as possible

Steps Toward Minimizing Bottlenecks

- Databasing can be a tedious chore.
- It is often helpful to design databasing “projects” (by taxa, locality, etc.) that are goals for the databaser as well as being useful sets of data that can be presented online.

The design of a new database system always has its detractors. It is rare for anyone today to recommend the use of file cards when presented with the capabilities of a computer database. But be prepared for those who favor pen and pad over computers when that database is unwieldy, behind schedule in its development, or has no support staff.