Preserving the Research Data of Natural History Collections

Geraldine Allen
Department of Biology
Archiving of research data is increasingly required by journals and funding agencies.

What are the advantages of archiving?

• Organization of the data for our own use
• Availability for other projects
Biodiversity-related research generates samples and specimens, along with their associated metadata.
Specimens documenting published research are required to be deposited in natural history collections (museums or herbaria).
Important uses of natural history collections:

• Reference specimens for ID of unknowns
• Documentation of occurrences (dates, places), including changes over time
• Documentation of biodiversity
• Sources of tissue samples for a range of analyses
New developments

• New research uses – specimens are used in ways not foreseen when they were collected (e.g., as sources of DNA).

• Digitization – information about biological collections (including images) is becoming available electronically.
About herbaria

A herbarium is a collection of dried plant specimens that are primarily used to describe and document new plant species and changes in the classification or naming of plants. Specimens housed in herbaria are used to better understand morphological variability among and between species, catalogue and identify the flora of a given region and to preserve a historical record of change in vegetation over time. Specimens are held in perpetuity and can last for hundreds of years if cared for properly.
Digitization of the UVic Herbarium

• Holdings: 50,000 plant specimens.
• Specimen label data captured for ~10,000 specimens (no images yet).
• Early data capture used a customized MS Access database (supported by UC Davis).
• In the last year, we have moved to a new Filemaker Pro database with much greater flexibility.
Advantages of the new database:

• Greatly improved error checking (e.g., detection of misspelled names, missing data fields, duplicated records).

• New uses for the specimen data (e.g., mapping, data extraction, exchange with other institutions) – can build these into the database design.
Next steps

• Regular archiving of database
• Imaging of specimens
• Contribution of specimen records to regional, national and eventually global biodiversity databases
**Taraxacum officinale (dandelion)**

[Map of Canada with points of interest]
Special-use collections

• The UVic Herbarium has a collection of ~10,000 leaf samples for DNA analysis (assembled in collaboration with RBCM).
• These are associated with herbarium specimens (at UVic, RBCM and elsewhere) and sometimes with extracted DNA.
• Metadata are stored in spreadsheets; no database yet (data archiving in early stages).
What have we learned from this project?

• Preparing data for archiving can be labour-intensive.
• No need to re-invent the wheel: many useful tools already exist.
What have we learned from this project?

Data archiving is worthwhile for many reasons.

• Compels researchers to get their data into good shape (saves time for further use)
• Opens up new uses and research applications (species distribution/ecological niche modelling)
• Promotes collaborations with others archiving similar data sets
Acknowledgements

Database development – Wendy Anthony

Collaborators (Royal BC Museum) – Ken Marr, Richard Hebda

Data entry – many work-study students and volunteers