

### Survey of Practices Related to Digitization of Specimen-Related Source Materials in Natural History Collections

iDigBio in collaboration with Yale Peabody Museum is planning a workshop focused on the digitization of specimen-related source materials in natural history collections. We define the term source materials to include field notebooks, catalogs, ledgers, cards, and other archival materials that contain specimen-related data. By digitize we mean the conversion of analog to digital, to include imaging, entering text data into an electronic database, converting analog sound recordings to digital recordings, etc. If you manage natural history collections, your answers to this survey will assist the planning team with (1) developing workshop content, (2) identifying those in the community who are involved in source material digitization, and (3) helping us create a profile of archival practices within natural history museums.

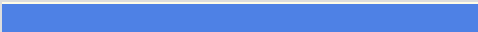
If you are interested in being contacted about the workshop, you will have an opportunity to provide your contact information. If you choose not to provide contact information, only your IP address will be registered; your identity and responses will be anonymous. Responses to this survey will be compiled and only composite results will be reported in any resulting presentations and publications.

Your participation in the survey is voluntary. There are no direct benefits or risks to you for participating, and no compensation. You may quit at any time or skip any item. You may withdraw your consent at any time without penalty.

The survey should take about 15 to 20 minutes to complete. Please respond by November 22.

If you have questions about this survey, contact Dr. Shari Ellis, Florida Museum of Natural History, University of Florida, 352-273-2066, [sellis@ufl.edu](mailto:sellis@ufl.edu). If you want more information about your rights as a research participant, contact the UF IRB Office, 352-392-0433, [irb2@ufl.edu](mailto:irb2@ufl.edu).

By checking the box below I acknowledge that I have read the information and agree to participate in this survey. If you do not wish to participate, please close your browser at this time.

|         |   |    |      |
|---------|---|----|------|
| I agree |  | 83 | 100% |
|---------|---|----|------|

My responses to the questions below about source documentation reflect my practices in the following collections. Check all that apply.

| Answer                    | Response | %   |
|---------------------------|----------|-----|
| Botany                    | 13       | 16% |
| Bryology                  | 2        | 2%  |
| Entomology                | 31       | 38% |
| Herpetology               | 16       | 20% |
| Ichthyology               | 13       | 16% |
| Invertebrate paleontology | 13       | 16% |
| Invertebrate zoology      | 15       | 18% |
| Lichenology               | 5        | 6%  |
| Malacology                | 10       | 12% |
| Mammalogy                 | 18       | 22% |
| Ornithology               | 17       | 21% |
| Paleobotany               | 7        | 9%  |
| Palynology                | 2        | 2%  |
| Phycology                 | 1        | 1%  |
| Vertebrate paleontology   | 16       | 20% |
| Mycology                  | 5        | 6%  |

Total Responses

82

3. My responses about source material digitization reflect my practices for the following preparation types. Check all that apply.

| Answer  | Response | %   |
|---|----------|-----|
| Flat sheets (e.g. herbarium sheets)             | 14       | 16% |
| Dry packets (e.g. bryophyte and lichen packets) | 7        | 8%  |
| Fluid-preserved vertebrates                     | 30       | 35% |
| Dry-preserved vertebrates                       | 31       | 36% |
| Fossils in trays and drawers                    | 22       | 26% |
| Shells in trays and drawers                     | 19       | 22% |
| Dried, pinned insects in trays or drawers       | 33       | 39% |
| Fluid-preserved invertebrates                   | 34       | 40% |
| Microscopic slides                              | 34       | 40% |
| Tissue  | 23       | 27% |
| Other. Please explain.                          | 12       | 14% |

| Other. Please explain.  |
|---|
| Observation   |
| Microfossil slides  |
| kodachrome slides, radiographs, audio tapes                                 |
| Paper documents   |
| stained & cleared whole   |
| field notes and maps  |
| archival material   |
| dried fungi in boxes, in-situ images photos or slides                       |
| Meteorites in trays, drawers, thin sections, powders, and oversized storage |
| Anthropology collections  |
| field notebooks, catalogs, cards, ledgers, etc.                             |
| catalog cards   |

4. Does your institution digitize source materials, to include field notebooks, catalogs, ledgers, cards, or similar original specimen-related source or archival materials?

| Answer | Response | %    |
|--------|----------|------|
| Yes    | 54       | 64%  |
| No     | 31       | 36%  |
| Total  | 85       | 100% |

5. If you would like to comment on why your institution does not digitize source materials, please do so in the box below.

#### Text Response

The reason is that our institution doesn't have many of those kinds of source materials; but we are planning on digitizing what we have in the near future.

There is currently no coordinated effort to digitize source materials. There are a few individual projects, ie. our library has digitized a few things with the "Field Book Project", but not for Botany. The field books of one of our historical collectors is also digitized, but there are many many more to do. Some of our collection managers have also expressed an interest in digitizing catalogues and ledgers.

No staff or time

We do keep the records, but not digitized. The actual fieldnotes are kept in a safe place. We lack resources to invent this big of a wheel. We would LOVE to participate in a best-practice, larger-community initiative. We wish these national/international efforts could include all NH collections rather than just biological materials.

Primarily due to insufficient resources (staff & time) and lower priority than the specimens (most are still not digitized).

no funding at this time

No facility to do so at the moment. We may try to have source materials digitized sometime in the future.

To my knowledge such source materials is not available in my collection.

Lack of time and human resources

Lack of that type of source material, lack of time, lack of funds

lack of staff resources to conduct such digitisation, although our collection is looking to employ a digitisation manager & so would look to doing this type of digitisation in the future ....

No funds.

We are presently planning to scan all catalogues and other important documents (19th Century) where fragile paper is used (true for vertebrates - mainly birds). On the other hand, digital cataloguing, at least for the last 20-30 years, is based on specimens, labels, old catalogues, travel books, field note books etc (for fish and herps but not for birds, based on the catalogue and not for mammals, based on labels only). Every department had its own strategy for how to digitize their Collections (natural history). Our experiences dates back to the late 1970s. For 18th Century stuff cataloguing is based entirely on secondary sources since all original labels were removed 1797.

We do not maintain a collection of such materials, with the exception of the non-lichenized fungi. This collection does maintain spore prints, written descriptions, and photographs. We are in the very beginning stages of digitization of the fungal collection.

Lack of resources... time or funding to spend much more time in the collection than basic preservation of it and occasional accession of new specimens. Our botanist is looking into digitizing the herbarium and is seeking external funding to do so.

So far, we have inadequate funding and manpower to make this a priority.

Time, staff, and money resources

We would like to, we just haven't gotten ourselves organized yet and we haven't secured any funding for it. Hoping to get it done in the next three years

We are in the process of applying for funding to purchase and support digital collection management software.

At present we do not have the ability or resources to do so.

Too much time required

It is not our first priority. At the present time we are absolutely devoted building databases with data from collections and disseminating their contents.

No funding do so - yet.

| Statistic       | Value |
|-----------------|-------|
| Total Responses | 24    |

6. What type of source material do you use to draft locality and/or collecting event descriptions?  
Check all that apply.

| Answer   | Response | %   |
|--|----------|-----|
| Specimen or collection object labels                     | 42       | 82% |
| Field notebooks/logs                                     | 46       | 90% |
| Catalogs/ledgers   | 37       | 73% |
| Specimen or collections object data cards                | 27       | 53% |
| Locality catalogs and/or cards                           | 15       | 29% |
| Standard data forms (e.g., geological measured sections) | 10       | 20% |
| Primary source literature materials and published works  | 18       | 35% |
| Annotated, hand-drawn, or other collector-generated maps | 20       | 39% |
| Field photographs  | 20       | 39% |
| Other. Please explain.                                   | 4        | 8%  |

Other. Please explain.

aerial or satellite imagery

GPS data

GPS and sensor data from ship's logs

digital maps

| Statistic       | Value |
|-----------------|-------|
| Total Responses | 51    |

7. For legacy specimen or objects being acquired today (e.g., donation), do you request available original source documentation to accompany these collections?

| Answer                         | Response | %    |
|--------------------------------|----------|------|
| Yes                            | 38       | 75%  |
| No                             | 3        | 6%   |
| It depends.<br>Please explain. | 10       | 20%  |
| Total                          | 51       | 100% |

It depends. Please explain.

If it is available and donor is willing

very rarely do we have donations to our collections, however, in the past with donations documentation has been requested

Documentation indicating that it was legally acquired, as needed, is required. Then any source documents.

if the material was collected pre-CITES or Lacey Act.

No: specimens acquired on exchange, from staff, other active collectors. Yes: if available from retired/deceased collectors

preferably if available from donor

If available

Specimens that do not have labels but are deemed of significant we would accept only with original source documentation.

Yes, but often no accompanying material

Only relevant permits for international material

8. For specimens or objects acquired through direct field work, what written or digital documentation do you request?

| Answer   | Response | %   |
|--|----------|-----|
| Field notes or logs                                      | 39       | 76% |
| Field labels   | 29       | 57% |
| Specimen or object label                                 | 36       | 71% |
| Maps   | 22       | 43% |
| Standard data forms (e.g., geological measured sections) | 11       | 22% |
| Images of locality and/or collecting event               | 18       | 35% |
| Georeference coordinates                                 | 41       | 80% |
| Other. Please explain.                                   | 10       | 20% |

Other. Please explain.

Any other data that might have been taken in the field or as part of analysis

we want to archive whatever the collector has

sometimes video

Spreadsheets/databases used to generate labels

ship's logs and sensor data (for deep sea collecting)

Accession paperwork - deed of gift and permits, specimen images of tissue vouchers

permits

Locality information for classified meteorites is maintained by the Meteoritical Society and

available via the online Meteoritical Bulletin Database

permits

spreadsheet of specimen and locality data

9. Are you digitizing analog source materials for new collections as they are acquired?

| Answer | Response | %    |
|--------|----------|------|
| Yes    | 30       | 60%  |
| No     | 20       | 40%  |
| Total  | 50       | 100% |

10. If you would like to explain your answer to the previous question, please do so in the space below.

**Text Response**

We scan all field notebooks used for current or past specimen collection. In addition, all historic specimen tags are stored in archival sleeves and data is transferred into the database. We hope to someday scan these as well.

We would like to, but do not have the funding.

As much as possible but specimens usually get cataloged first  
our specimen data are entered into electronic databases

As we are able - there is such a backlog of items from our collection that the amount of data from new collections tends to have more detail (GPS coordinates, digital pictures, etc.) that the rush is to digitize the older items on VHS, black/white paper photos, slides, & lots of notes, drawings etc on paper.

The "real answer" is yes/no: - yes - the specimen labels if not already provided in digital form.  
- no - other ancillary material; we receive very little that contains info beyond the specimen unless it is a donation of someone who is no longer active in the field.

not enough resources and digitizing process not established.

Generally we only get coordinates or a simple description (though we would like more). What we do get we scan as pdfs and attach to the relevant locality record.

Generally this means typing data from newly acquired specimens and their associated labels/logs into our database. We don't currently have a way to upload digital information (e.g. GPS coordinates, sensor logs). We are generally not imaging them.

Usually in the form of field note data

New plant specimens are imaged and databased from specimen labels. New fungal specimens are databased from specimen labels, specimen cards with detailed notes on features including microscopy are scanned, and any images submitted are uploaded with the specimen.

We do not have the resources to do this.

Our collection was established recently with a very large legacy collection of 50,000 lots of slides and fluid preserved material, and we are in the process of digitizing the handwritten data ledger and specimen labels into an online database. Data from new material coming into the collection is immediately entered into this database.

We digitally catalogue the specimens with the fullest provenance data available. But digitizing the source materials is not typical or uniform. Some field notes are digitized, most are not.

| Statistic       | Value |
|-----------------|-------|
| Total Responses | 14    |

### 11. Do you Image?

| Question                             | Yes | No | Total Responses |
|--------------------------------------|-----|----|-----------------|
| Specimen or collection object labels | 28  | 19 | 47              |
| Field notebooks/logs                 | 33  | 15 | 48              |
| Maps                                 | 16  | 28 | 44              |
| Catalogs/ledgers                     | 26  | 21 | 47              |
| Photo media                          | 31  | 14 | 45              |
| Catalog cards                        | 20  | 25 | 45              |
| Locality catalogs and/or cards       | 10  | 29 | 39              |
| Standard data forms                  | 13  | 24 | 37              |

### 12. Level of Priority

| Question                             | High priority | Medium Priority | Low Priority | Total Responses |
|--------------------------------------|---------------|-----------------|--------------|-----------------|
| Specimen or collection object labels | 14            | 13              | 12           | 39              |
| Field notebooks/logs                 | 17            | 11              | 13           | 41              |
| Maps                                 | 5             | 7               | 16           | 28              |
| Catalogs/ledgers                     | 13            | 10              | 14           | 37              |
| Photo media                          | 7             | 17              | 12           | 36              |
| Catalog cards                        | 10            | 7               | 15           | 32              |
| Locality catalogs and/or cards       | 5             | 5               | 14           | 24              |
| Standard data forms                  | 6             | 6               | 11           | 23              |

### 13. Do you transcribe from physical objects?

| Question                             | Yes | No | Total Responses |
|--------------------------------------|-----|----|-----------------|
| Specimen or collection object labels | 40  | 4  | 44              |
| Field notebooks/logs                 | 32  | 13 | 45              |
| Maps                                 | 16  | 27 | 43              |
| Catalogs/ledgers                     | 27  | 16 | 43              |
| Photo media                          | 15  | 27 | 42              |
| Catalog cards                        | 21  | 21 | 42              |
| Locality catalogs and/or cards       | 12  | 29 | 41              |
| Standard data forms                  | 14  | 25 | 39              |

### 14. Level of Priority

| Question                             | High priority | Medium Priority | Low Priority | Total Responses |
|--------------------------------------|---------------|-----------------|--------------|-----------------|
| Specimen or collection object labels | 35            | 4               | 2            | 41              |
| Field notebooks/logs                 | 21            | 7               | 10           | 38              |
| Maps                                 | 5             | 8               | 17           | 30              |
| Catalogs/ledgers                     | 21            | 4               | 11           | 36              |
| Photo media                          | 6             | 6               | 17           | 29              |
| Catalog cards                        | 17            | 2               | 13           | 32              |
| Locality catalogs and/or cards       | 9             | 1               | 16           | 26              |
| Standard data forms                  | 10            | 4               | 12           | 26              |



15. Do you transcribe from images?

| Question                             | Yes | No | Total Responses |
|--------------------------------------|-----|----|-----------------|
| Specimen or collection object labels | 14  | 31 | 45              |
| Field notebooks/logs                 | 13  | 31 | 44              |
| Catalogs/ledgers                     | 8   | 35 | 43              |
| Catalog cards                        | 6   | 37 | 43              |
| Locality catalogs and/or cards       | 5   | 36 | 41              |
| Standardized data forms              | 5   | 36 | 41              |
| Photo media                          | 7   | 35 | 42              |
| Maps                                 | 6   | 36 | 42              |

16. Level of Priority

| Question                             | High priority | Medium Priority | Low Priority | Total Responses |
|--------------------------------------|---------------|-----------------|--------------|-----------------|
| Specimen or collection object labels | 8             | 10              | 8            | 26              |
| Field notebooks/logs                 | 11            | 4               | 9            | 24              |
| Catalogs/ledgers                     | 7             | 5               | 8            | 20              |
| Catalog cards                        | 5             | 4               | 8            | 17              |
| Locality catalogs and/or cards       | 2             | 3               | 11           | 16              |
| Standardized data forms              | 4             | 2               | 11           | 17              |
| Photo media                          | 1             | 7               | 13           | 21              |
| Maps                                 | 3             | 4               | 11           | 18              |

17. Are you accepting born digital source materials?

| Answer | Response | %    |
|--------|----------|------|
| Yes    | 27       | 61%  |
| No     | 17       | 39%  |
| Total  | 44       | 100% |

18. Do you digitize from analog audio data?

| Answer | Response | %    |
|--------|----------|------|
| Yes    | 4        | 9%   |
| No     | 42       | 91%  |
| Total  | 46       | 100% |

19. What challenges have you encountered transcribing data from physical objects of source materials?

#### Text Response

old labels are damaged and no longer readable

Handwriting

Reading and understand hand-written text.

Challenges have included deciphering handwriting, detection and correction of errors, handling fragile source materials (e.g. records/specimens date from 1840s), and managing time/staffing resources (i.e. there is always a backlog).

Loss of field notebooks and data over the years and disassociation of data cards from objects.

Difficulty interpreting handwriting and abbreviations; non-standardized locations; label field not matching database fields; having to move specimens to get to label or to view all portions of label; 1 field note book page containing information about multiple specimens and/or localities; how to handle things like locality sketched and cross sections

Limited staff and volunteer time. Minimal volunteer experience and expertise.

Sheer volume

we have scanned field notebooks and assembled files into .pdf format. The process is cumbersome and time consuming. Imaging with a camera may be a faster solution.

Difficult to read hand-writing, time intensive, sometimes field # notebook doesn't match the field # on the specimen

Bad handwriting and horrid grammar.

We are transcribing from physical objects vs. images, as we have a limited number of computers to use. Many items that need to be transcribed & changed come from little slips of paper in specimen boxes. Notes are routinely found everywhere, catalogs, put in office library books, taped on contact lists etc. Found a drawing of a specimen as it was found in the field from who knows where or when but we are narrowing it down. This seems to be a common place with not only our collection. We have decided it was more important to upload actual details into the computer before we spend time scanning & then uploading. The goal is to prepare items for a database like Specify from a customized Access database that I literally threw together on the fly from individual excel spreadsheets.

Interpreting handwriting, fading ink, uncommon abbreviations, localities that no longer exist, bad spelling are the most common issues we've encountered.

Proofing, id's, and illegible handwriting.

Finding the time to do any these tasks requires more staff, or supervision of students. Money to pay for these projects is necessary, but not forthcoming. Making curators see the benefit if digitization of materials is also a challenge.

Staff time - funding

reading hand-written data reading faded data

Physical handling challenges for dry pinned insects, and for specimens in wet storage (eg. ethanol/formalin),

We simply lack the funding to pay for personnel to do this.

Accurately reading handwriting, esp. on older specimens in non-English language.

None

Transcription errors between two sources (e.g., label vs catalogue book). Not enough data (missing georeference of enough details to obtain reasonable georeference).

Handwriting is our biggest challenge (plus the occasional lost or misplaced label).

Conflicting data from physical objects that should be referring to the same items. Loss of legibility for older carbon copy forms when only the carbon copy remains.

Bad handwriting, unclear abbreviations, obsolete names (taxon and place), foreign languages, inadequate original data, issues with how much data to enter vs how long it will take (for data

beyond the basics, such as measurements, habitat), inadequate QA/QC and ways to catch errors during the entry process, lack of staff to carry out and oversee work

Incomplete data, illegible data and data standard issues

Illegible handwriting, incomplete information, time constraints, duplication from material scattered through collections

Loss of data through source material damage. Questions about credibility of data. Handwriting legibility

We often have three or four labels with a single specimen. The labels can conflict and/or have different levels of data. Difficult to train interns on how to interpret this. The labels are often old, curled, and handwritten. So can be hard to read and time consuming. Some labels written on napkins, receipts, cardboard ect. Hard to read. Some labels are on front of drawer or on specimen themselves. We do not have a lot of maps, measured sections or fieldnote books. I handle these myself when they occur.

lack of resources; student interns have trouble reading historic handwriting

Funds to support data entry.

Fragility of material

preservative shrinks specimens and distorts shapes

reading hand-writing can be difficult. Training employees to recognize what the words could be, in the context of the specimen, is critical. Also, the labels are often wet or possibly dirty from the specimen (blood, etc.) and must be treated appropriately (gloves, etc.). This would also be the case with imaging these labels.

Difficulties in reading handwriting from legacy material and determining localities when only a town is given.

Legibility, interpretation of handwritten abbreviations, deterioration, risk of damage from handling

Handwriting that is difficult to read; out of country localities can be difficult to understand/decipher; sometimes specimen labels and the corresponding ledger do not match up – which one is the right one?; localities can be difficult to decipher when collectors use local names for an area; abbreviations can be very difficult to figure out; dates – does 1/11/87 = 11 January 1987 or 1 November 1987?

Inability to read handwriting. Discoloration or damage to card due to storage directly with specimens.

| Statistic       | Value |
|-----------------|-------|
| Total Responses | 38    |

20. What challenges have you encountered transcribing data from images of source materials?

Text Response

Handwriting; confusion amongst file names

Locating the hardware to digitize from microfilm.

We presently do not transcribe data from images of source materials, but are interested in do so in the future.

None. Not many images of source materials to work with as of yet.

Difficulty interpreting handwriting and abbreviations, though imaging software tricks make it easier (increasing size, contrast, etc.); non-standardized locations; label field not matching database fields; 1 field note book page containing information about multiple specimens and/or localities; how to handle things like locality sketched and cross sections

Sheer volume is too much to handle, backlog of images and little time to work through collection.

n/a

we don't do this because the specimen label is generally under the specimen so too hard to read

Penciled notes do not image well.

It would be very difficult to transcribe the majority of our data from images of source materials - the effort it would take to scan items so that they could be readable would take more time than utilizing the physical object. Handwriting can be difficult to read & written lightly in pencil.

Occasionally we have had to scan & darken source materials just to read them. Once items have been cataloged we plan on scanning in all documents, but the biggest push is to get all specimens into a database with all of the corresponding locality, field data etc, as cleaning up the data is going to be a huge undertaking just to be able to upload the details into a database like Specify. (Specify doesn't really like odd bone from N. of the pink siltstone near .....

Unclear/blurry images, image filenames mislabeled, not all of object/data source image captured in image...

Proofing, id's, and illegible handwriting.

Staff time - funding

OCR sucks despite what some are reporting in order to obtain funding. It's faster and more efficient to just keystroke the data into the appropriate fields after reading the image text.

However, we are developing a voice data control center to transcribe image data, which will make the process go much faster with less errors.

Issues with accuracy of transcribed data when transcribed by persons unfamiliar with the collections, issues with discrimination between essential and non-essential data categories, issues with poorly curated material (errors in identification are perpetuated, or specimens cannot be reliably tagged in the data)

We simply lack the funding to pay for personnel to do this.

none

nothing additional to those mentioned above for Physical Objects.

We haven't done this.

Not applicable as we transcribe from the original materials. We make images of materials after recovering priority data.

Lack of an efficient way to image data cards, lack of a way to show data card image next to data entry screen (i.e. no integrated workflow), lack of staff with appropriate knowledge. We would love to have a way to harness people out in internet space to help with this but much too technically complicated for us and insufficient staff to oversee.

Same as above

quality of image issues, naming of files

handwriting legibility

N/A

We often have three or four labels with a single specimen. The labels can conflict and/or have different levels of data. Difficult to train interns on how to interpret this. Digitizing the label makes them easier to read and nice to see all 4 labels at once.

n/a

lack of resources

Funds to support the work.

scanning works well with dark ink and typewritten information but penciled in coordinates are often washed out.

poor focus, the ink may not be legible at that lighting, sometimes it is not the ink that I read, but the embossment left by the pen, photographing at the right angle is hard.

Legibility, interpretation of handwritten abbreviations, deterioration - even more so than above

We take photos of the physical objects and locality cards in the same shot. Sometimes the locality card and/or locality number on the specimen cannot be confirmed due to the angle of the shot or because the specimen, not the card, is in crisp focus.

| Statistic       | Value |
|-----------------|-------|
| Total Responses | 33    |

21. Do you use interpretations of source materials to enhance specimen data (e.g., would you change Persia to Iraq to enforce the use of modern terminology)?

| Answer | Response | %    |
|--------|----------|------|
| Yes    | 31       | 70%  |
| No     | 13       | 30%  |
| Total  | 44       | 100% |

22. Please explain how you use interpretations to enhance specimen data or provide an example.

#### Text Response

Outdated place names are updated, if known.

Updating geopolitical names of countries.

Updated names of countries are changed from original source records upon entry into our specimen database (e.g. Tanganyika would be changed to Tanzania). The verbatim locality description is also text entered into the database.

We try our best to interpret to the original description and add the present day equivalent.

In addition to a verbatim locality we parse our data into separate fields so that it can be plotted via electronic mapping; we are also contemplating having a field for name of country at collection to handle the Persia component of the example you provided us in addition to the country field that handles the Iraq component of the example you provided us

Conversion to modern geographic and stratigraphic terminology, i.e. Persia to Iran

Expanded from only entering legal descriptions of collections sites on labels to include GPS locations or GIS shapefile with location site.

Our geographic name data are updated (enhanced) regularly as are our taxonomic data. Other enhancements such as georeference data are stored in associated / related files.

We would make corrections for new political names but always annotate in the notes what the original said

Currently we do not have to use interpretations, however, we have decided that we would use current terminology while also preserving the original data in a notes field. As we are able to clean up our field locality data for Madagascar, this will be very important.

If an old specimen is cataloged, we would indicate the changed locality names when appropriate.

update geopolitical names to reflect current usage

Routine interpretations include expanding abbreviations, using full names of collectors and other parties, using contemporary place names adjacent to historic names, including contemporary taxonomy against old names.

We would change geography/provenance to modern equivalents. We include verbatim data in some fields (e.g., date) but also standardize so as to be searchable.

Provide new name but always keep original term in record.

Somewhat dependent on the collection, but typically will change country name to current country name if possible, often will add county/province name and update this to reflect current geography (e.g. splits in counties/provinces since collection date); also add continent/ocean which is almost never on the tag; we will also use non-countries as the country where relevant for biological purposes (e.g. Guam is the country even though it is a US territory). Also typically update taxonomy to reflect current knowledge (generally following a specific authority). Will also annotate when we find errors on tags (typically related to place). May also add collector name if we know what it is even though it isn't on the tag/label.

Geographic names and taxonomic names all updated to latest name, no abbreviations in text strings, standardization of text strings (locality names etc.)

Legacy geographic names are converted to modern

Will add to locality descriptions for clarification based on GPS data. We will place secondary descriptions in brackets [] to keep our descriptions separate from original locality descriptions.

I will record the data as it is presented on the specimen ID label, but in another field of our database I will note the location's current name for clarification.

historic cataloged information is kept in verbatim format but a modern interpretation of location or updated scientific names is added to the records

Interpretations as given in the example would be used for ease in searching, the original

information would be stored in a notes field so the information is not lost.

Locality names have changed over time (Bavaria, Persia, etc.), as have classification schemes. While original terminology and materials are preserved, modern terminology is used in main database fields.

We indicate both terms... the original and then in [] the new term. both terms are then searchable and the old geographic term manages to retain the historical context.

This is for mapping purposes and to show distributions; information that has been interpreted is kept in the notes field.

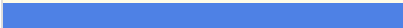

The original data are captured as a "verbatim" entry, but locality, date, collector etc. are standardized in a separate data columns. For example, a locality written variously as "Anchorage, Alaska," "Anch., AK", or "Anch." would be entered as such in "verbatim locality" but also entered separately as "Anchorage" in a standardized specific locality field embedded within higher geography. A species identification on a label based on outdated taxonomy would be captured as recorded - but then entered separately in a controlled taxonomy field with a currently accepted name.

Any enhancements made to the original data are within brackets [].

Some county names have changed over time. The county name would be updated in our database to reflect current terminology so that the georeferencing visible to the public (which is done to nearest county) is correct.

| Statistic       | Value |
|-----------------|-------|
| Total Responses | 28    |

23. Do you use source materials to improve the precision or accuracy of data?

|   | Response | %    |
|---|----------|------|
|  | 37       | 84%  |
|  | 7        | 16%  |
|   | 44       | 100% |



24. Please explain how you use source materials to improve accuracy of data or provide an example.

#### Text Response

Field notebooks and specimen provide a wide variety of data that was not written in the original handwritten catalog for our museum, such as morphometric measurements, behavior, coloring, etc.

We refer back to field notebooks to improve accuracy.

For instance, if a discrepancy is found, typically we use other source materials recorded by a more knowledgeable faculty member or staff on the same day to correct or confirm the error.

If a collector drew a circle on a map we are using that radius as our error, in addition we are using field note books to obtain information from cross sections or locality sketches to help narrow where accuracy is

Georeference based upon most specific locality info available.

Field notes are used to amend or correct electronic catalogue entries.

We check every source possible to confirm data

Ex: Specimen located, card blank except for a catalog & field number. Go to catalog - number in catalog has never been used (as far as we can tell - in the middle of blanks), go to log book - field number in log book with no details & no catalog number. Go to field book (yes sadly 3 books) and in the field book is a tiny little note on the side where specimen was sent to France on loan. Go loan folders (not in any order) & find nothing - well details were written on a scrap piece of paper in an envelope found in a book of misc. information. Further details were able to be entered on specimen card, all books & in database while preparing a logical order when scanning of paperwork can begin.

We use collector notes to verify specimen tags; use geographic sources and google earth to add in state, province, or municipality information; oral interviews of collector to clarify problems/discrepancies...

Georeferencing was done through MaNIS, ORNIS & HerpNet, but the locations that were in the database can usually be improved (~0%) by looking at collecting catalogs for additional location data.

If a locality is georeferenced, we would use whatever tools necessary to achieve a satisfactory result. If there is a question on the locality, we may go to the field notebook to see if there is more information that would provide better resolution of the locality or date.

Mapping software to check georeference data.

use written descriptions of physical localities to try to reduce the uncertainty

We use external checklists or faunal directories to obtain the most current taxonomic content, or recent publications when large groups have been revised.

Use of online atlases/gazeteers to try to locate place names.

Use original GPS data

Compare specimen label with catalogue record to make sure no info has been dropped on label; consult field data to add anything relevant or additional precision.

Specimen labels may only have general locality data, but if they have a date or a collection number, we go back to the original field notes (if we have them), to fill out more information about that locality.

We use both the original wording but also translate into modern language (as in the question example with Persia to Iraq).

Refer back to field catalogs to help deal with handwriting issues or if we suspect an error on the label. Also use source material in cases where the label/tag has incomplete/abbreviated data. Also helpful if tags/labels are damaged. Also to deal with past situations where data may have been entered from the field notes but not transcribed onto the specimen label. It is also sometimes easier to correct legacy data from the field notes rather than pulling all the



specimens involved. Field notes can also help with general context information when starting work on a new collection.

Field notes, maps etc from existing collections are used to check accuracy of data in the database

Field notes make it more clear when georeferencing (miles by car or straight line) and cleaning data (labels reading June 31...)

We verify data entered digitally with the original fieldnotes and specimen labels. Often find transcription errors and update our digital databases to more accurately represent the original data sources.

For the first question we would keep Persia but also have a modern standard, Iraq. If there is additional information in the field book, accessions record or strat section that is relevant to the specimen we would add that to the specimen data. i.e If the label listed "South of Racine" but the a field map shows the location at the outcrop on highway 20, we would add highway 20 outcrop to the precise locality.

we consult historic expedition reports, for example, for clarification of material collection data

Source materials are used to help with georeferencing and confirmation of identification for fungi in which in-situ images will provide additional characters than the specimen alone.

In some cases, early published papers provide much more detail on collection circumstances.

Photos can indicate the physical state of a specimen at a given time (particularly useful if a specimen has been sectioned or cut, since).

Use descriptions of locations to narrow down geographic areas

Always check the field data to make sure they were there on that date, etc.

We compare captured ledger data with original specimen labels, and also check these records against published literature.

A published paper might contain more specific locality data than what can be found on the specimen label.

Some of the transcriptions of locality data do not include the locality descriptions. By returning to the ledgers and digitizing these locality descriptions it is often possible to georeference localities more accurately.

| Statistic       | Value |
|-----------------|-------|
| Total Responses | 32    |

25. If you interpret or change data as it is transcribed, do you cite the interpretation and the source of the change?

| Answer | Response | %    |
|--------|----------|------|
| Yes    | 32       | 74%  |
| No     | 11       | 26%  |
| Total  | 43       | 100% |

26. Please explain or provide an example of how changes are made to data as it is transcribed.

#### Text Response

Country name changes are made following an on-line source for standardized country names. In digital format information is simply changed. However on the written specimen card a pencil line is drawn through the incorrect information and correct information is written in or perhaps a new card is made.

A notes field incorporates information on source or sources of information

If species identification is incorrect or old names are used (on herbarium labels specifically), inserts are added to reflect the name changes.

Precise localities are commonly "pigeon holed" into higher geographic categories (country state, county) as needed.

Verbatim data are entered and then followed by annotations of changes with appropriate sources and the name and date of the person making the change.

New information on specimen entered into appropriate field - in notes field: original data, Researcher ID of new data, date changed, database change date & initials of database editor.

Specimen identification changes - noting the person, date, and reason why a specimen's ID was changed; in tissue collections, if any number could not be read on the vial or if there is some related discrepancy, the problem is noted, dated, and initialed.

Georeferenced locations are annotated when more precise or additional data are added, with corrected from ... to ... initials and date.

note specific reference to the source of the data used to make the change(s)

eg. "NNE of Bald Nob FR, Bellangry NP" would become "NNE of Bald Nob Flora Reserve, Bellangry National Park" Note that we do not always expand abbreviated data where the meaning is clear, but for abbreviations like FR which could be Flora or Fauna Reserve we use interpretations to make things clearer.

If we have the personnel time we will check data for accuracy, ie. map coordinates.

We have data fields for "verbatim entry" - often used for dates, sometimes for locality (depending on project). - we also use a "specimen notes" field where most, but not necessarily all, such changes are recorded.

provide source of document or map or person providing information.

Some places are misplaced in the original material or on the label. We preserve this and then add a correction within square brackets.

all original data are retained and changes documented

Any changes we make to locality data we explain in a section of our database called "Locality Notes". We also enter the date and the name of who made the change. If we add anything to the original locality description we mark it with brackets.

We would include an image of the original label and an image of any other source that we got the data from.

updated locations or scientific names are logged with date and name of individual making change

Any changes that are made are to help conform to common searches or provide up-to-date reference points for georeferencing. For example, legacy fungi specimens often do not have the full name of the collector on the specimen label only initials. In the collector field, we would write out the full name of the collector if know. Or, if a label misspells a county name, we would fix that name for the county field.

Modern country names are used rather than those from, for example, the 1800s. A scanned copy of the original specimen documentation is directly linked to each database entry. Changes in physical status are noted (if thin sections are made, the specimen mass changes, for example).

Coordinates will be entered into the database in the format in which they were collected, but will be converted to decimal degrees for use on labels.

Dates are almost never written correctly so they must be formatted correctly for our collection database. Names of people may be shortened or initials only used. Yet, if we know that some set of initials is a specific collector, we enter it by name, not initials. Know the collectors and know the time-period and all the short-cuts they took in field notebook or on specimen labels will start to make sense.

NY would be New York; 1/1980 would be January 1980

As described earlier. Abbreviations are interpreted, but annotated as to who made the interpretation. The actual image of the ledger will be linked with the specimen record for visual confirmation of the transcription.

Locality: Pamlico River, [at Mauls Point, ca. 9.8 air miles SE center Washington], [see remarks]

Locality remarks:[Original label does not indicate exact location on Pamlico River, but gives station # of MP-3. Examining similarly prefixed station numbers (MP-\*), however, indicates "Maul's Point".]

| Statistic       | Value |
|-----------------|-------|
| Total Responses | 26    |

27. On your website, do you have links to source material from your institution and/or others?

| Answer | Response | %    |
|--------|----------|------|
| Yes    | 10       | 24%  |
| No     | 32       | 76%  |
| Total  | 42       | 100% |

28. Do you use any standards, best practices, or technologies to make your digitized source materials more useful/referenceable?

| Answer | Response | %    |
|--------|----------|------|
| Yes    | 27       | 64%  |
| No     | 15       | 36%  |
| Total  | 42       | 100% |

29. Which of the following standards, best practices, or technologies do you use? Check all that apply.

| Answer  | Response | %   |
|---|----------|-----|
| Standardized metadata for each source                 | 15       | 60% |
| Standardized filenames for each source                | 16       | 64% |
| Global unique identifiers associated with each source | 10       | 40% |
| Optical Character Recognition (OCR)                   | 7        | 28% |
| Parsing tools   | 6        | 24% |
| Markup tools  | 2        | 8%  |
| Other. Please explain.                                | 3        | 12% |

Other. Please explain.

Manis/HerpNET/Ornis Standards for Georeferencing

Finding aids that describe the source materials.

Directly linked to database entry for applicable specimen, so all information can be viewed simultaneously

30. If you are potentially interested in participating in a workshop on digitization of source materials or anticipate that someone at your institution might be, please provide your name, e-mail address, and institution name in the spaces provided so that we can contact you.

| Name                 | Email Address               | Institution   |
|----------------------|-----------------------------|---|
| Jessa Watters        | jwatters@ou.edu             | Sam Noble Museum, Norman, OK                          |
| Dr. Debra Trock      | dtrock@calacademy.org       | California Academy of Sciences                        |
| Anielkoemar Gangadin | anilkumar.gangadin@uvs.edu  | National Zoological Collection of Suriname (NZCS)     |
| Ilyn sharp           | llyn@vt.edu                 | Virginia Tech Geosciences                             |
| Susan Borkin         | borkin@mpm.edu              | Milwaukee Public Museum                               |
| Chris Phillips       | caphilli@illinois.edu       | IL Natural History Survey                             |
| Laura Abraczinskas   | abraczi1@msu.edu            | Michigan State University Museum                      |
| Ashley Quinn         | ashley.quinn@gcsu.edu       | Georgia College Natural History Museum                |
| Margaret Landis      | paleocatstar@ou.edu         | SNOMNH  |
| Glenn Storrs         | gstorrs@cincymuseum.org     | Cincinnati Museum Center                              |
| Elizabeth Hill       | hilleliz@grinnell.edu       | Grinnell College (Conard Environmental Research Area) |
| Greg Schneider       | ges@umich.edu               | UMMZ  |
| Christy Bills        | cbills@nhmu.utah.edu        | Natural History Museum of Utah                        |
| Pat Holroyd          | pholroyd@berkeley.edu       | University of California                              |
| Tim Whitfeld         | Timothy_Whitfeld@brown.edu  | Brown University                                      |
| L. Catherine Riddle  | criddle@duke.edu            | Duke Lemur Center                                     |
| Heath Garner         | heath.garner@ttu.edu        | Division of Fossil Primates                           |
| Mark O'Brien         | mfbrien@umich.edu           | Museum of Texas Tech University                       |
| Ryan Caesar          | rmcaesar@hawaii.edu         | Univ. of Michigan Museum of Zoology                   |
| Vince Eckhart        | eckhart@grinnell.edu        | University of Hawaii                                  |
| Nicole Fisher        | nicole.fisher@csiro.au      | Grinnell College                                      |
| Andrew Miller        | amiller@inhs.illinois.edu   | ANIC  |
| Jack Schuster        | jschuste@uvg.edu.gt         | University of Illinois Urbana-Champaign               |
| Ron Eng              | rceng@uw.edu                | Universidad del Valle de Guatemala                    |
| Dave Britton         | dave.britton@austmus.gov.au | Burke Museum  |
| Lynn Kimsey          | lskimsey@ucdavis.edu        | Australian Museum, Sydney                             |
| Rich Rabeler         | rabeler@umich.edu           | University of California, Davis                       |
| Jean-Marc Gagnon     | jmgagnon@mus-nature.ca      | Univ of Michigan Herbarium                            |
| Jean Woods           | jwoods@delmnh.org           | Canadian Museum of Nature                             |
|                      |                             | Delaware Museum of Natural History                    |

|                      |                                  |   |
|----------------------|----------------------------------|---|
| Andy Bentley         | abentley@ku.edu                  | KU Biodiversity Institute   |
| Paul Sweet           | sweet@amnh.org                   | AMNH  |
| Molly Phillips       | Molly.Phillips@MyFWC.com         | Fish and Wildlife Research<br>Institute Ichthyology<br>Collection                           |
| Joe Whittaker        | jwhittak@cord.edu                | Concordia College   |
| Paul Mayer           | pmayer@fieldmuseum.org           | The Field Museum  |
| Joe McHugh           | mchugh.jv@gmail.com              | University of Georgia   |
| cindy opitz          | cindy-opitz@uiowa.edu            | Museum of Natural History,<br>The University of Iowa  |
| Melissa Islam        | melissa.islam@botanicgardens.org | Denver Botanic Gardens  |
| Rebekah Hines        | rrhines@asu.edu                  | Arizona State University  |
| Joan Herrera         | joan.herrera@myfwc.com           | Florida Fish and Wildlife<br>Research Institute   |
| Greg Watkins-Colwell | gregory.watkins-colwell@yale.edu | Yale Peabody Museum   |
| Michelle Knoll       | mknoll@nhmu.utah.edu             | Natural History Museum of<br>Utah   |
| Maribel Campbell     | campbell@carachupa.org           | Division of Parasitology,<br>Museum of Southwestern<br>Biology, University of New<br>Mexico |
| Kirsten R Bophy      | kbrophy@stamfordmuseum.org       | Stamford Museum & Nature<br>Center  |
| Irit Zohar           | zoharir@gmail.com                | Oranim College, Israel  |
| Erica Clites         | eclites@berkeley.edu             | University of California<br>Museum of Paleontology  |

| Statistic       | Value |
|-----------------|-------|
| Total Responses | 45    |

### 31. Gender


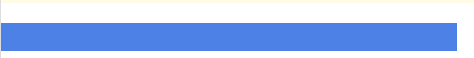
| Answer | Response | %    |
|--------|----------|------|
| Female | 36       | 52%  |
| Male   | 33       | 48%  |
| Total  | 69       | 100% |

### 32. Disability. Check all that apply.



| Answer                         | Response | %   |
|--------------------------------|----------|-----|
| Hearing Impairment             | 1        | 2%  |
| Visual Impairment              | 1        | 2%  |
| Mobility/Orthopedic Impairment | 0        | 0%  |
| Other (Enter description)      | 0        | 0%  |
| None                           | 59       | 97% |

Other (Enter description)




### 33. Ethnicity

| Answer                 |   | Response | %    |
|------------------------|---|----------|------|
| Hispanic or Latino     |  | 3        | 4%   |
| Not Hispanic or Latino |  | 65       | 96%  |
| Total                  |   | 68       | 100% |

### 34. Race definitions. You may check more than one.

| Answer                                    |   | Response | %   |
|---|---|----------|-----|
| American Indian or Alaskan Native         |   | 0        | 0%  |
| Asian                                     |    | 3        | 5%  |
| Black or African American                 |   | 0        | 0%  |
| Native Hawaiian or Other Pacific Islander |   | 0        | 0%  |
| White                                     |  | 63       | 95% |

### 35. Citizenship

| Answer                 |   | Response | %    |
|------------------------|---|----------|------|
| U.S. Citizen           |  | 56       | 81%  |
| Permanent Resident     |  | 4        | 6%   |
| Other non-U.S. Citizen |  | 9        | 13%  |
| Do not wish to provide |   | 0        | 0%   |
| Total                  |   | 69       | 100% |