Database Field Guide

Fields denoted with (*) are required and may not be left blank. Those not denoted with this may be assumed to allow empty values, disregarding the specified formatting for these cases where the data for the field is not present.

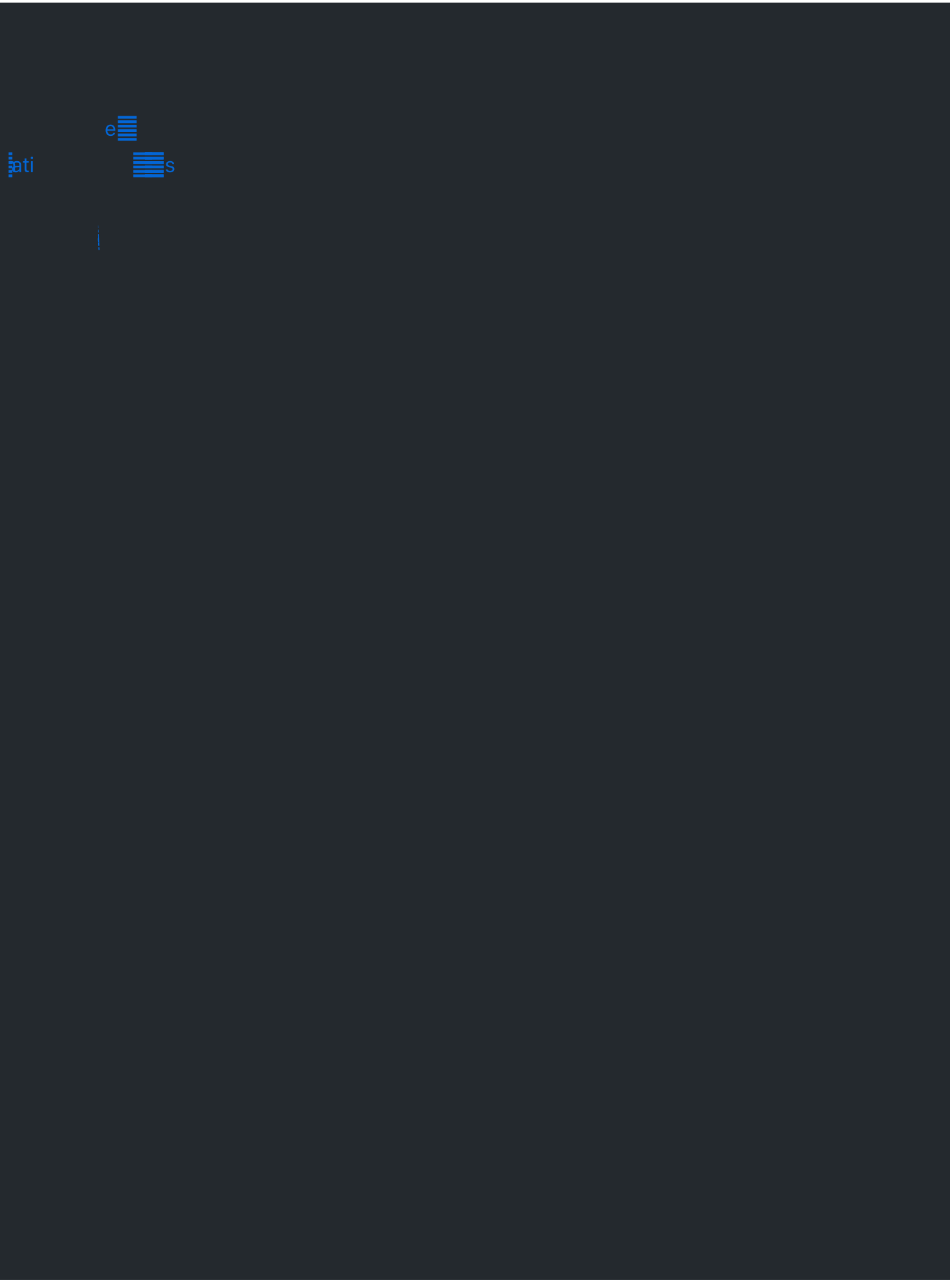
Fields that do not have any generalized, restricted formatting will have relaxed insertion requirements. You may see additional information about what the field is by referencing the Darwin Core guide.

Note: Some fields are automatically generated/determined, these fields should **not** be included in the CSV for batch insertions. An example is `recordEnteredBy`, the credentials of the logged in user will be applied for this information.

Available Specimen Fields:

(click to view formatting information)

- catalogNumber
- otherCatalogNumber
- recordNumber
- otherIdentifier
- projectNumber
- order_
- superfamily
- family
- subfamily
- tribe
- genus
- subgenus
- specificEpithet
- infraspecificEpithet
- identificationQualifier
- recordedBy
- otherCollectors
- identifiedBy
- dateIdentified



- associatedSequences
- associatedReferences
- withholdData
- reared
- recordEnteredBy
- modifiedInfo
- fieldNotes

catalogNumber *

This is the LEP number. The general format is:

```
LEP followed by 5-8 digits.
No punctuation allowed.
Case sensitive.
Empty values NOT accepted (i.e. must be included)
```

Pa ing Example

```
LEP12345  LEP123456  LEP1234567  LEP12345678
```

Failing Example

```
LEP1234  LEP-123456  lep1234567
```

otherCatalogNumber

This is the MGCL number. The general format is:

```
MGCL_ followed by 6-8 digits.
No punctuation allowed.
Case sensitive.
```

Pa ing Example

```
MGCL_123456  MGCL_1234567  MGCL_12345678
```

Failing Example

```
MGCL_12345  MGCL_123456789  mgcl_12345678  MGCL1234678  MGCL-1234678
```

recordNumber

This is a Collector's record number. There is no specified, mandatory formatting for this field.

Pa ing Example

```
LK45
```

otherIdentifier

This is a new field and has not had any formatting guidelines generated yet.

projectNumber

This is a new field and has not had any formatting guidelines generated yet.

order_

Linnaean Order. The formatting is:

```
Capitalized first letter for each separate word.
No random capitalization (i.e. no capitalizations in the middle of a word).
No punctuation.
```

Pa ing Example

```
Order  Order Name
```

Failing Example

```
order  order name  order-name  OrDeR  Order?
```

superfamily

Linnaean Superfamily. The formatting is:

```
Capitalized first letter for each separate word.
No random capitalization (i.e. no capitalizations in the middle of a word).
No punctuation.
```

Pa ing Example

```
Superfamily  Super Family
```

Failing Example

```
superfamily  super family  super-family  SuPerFamilY  Superfamily?
```

family

Linnaean Family. The formatting is:

```
    Capitalized first letter for each separate word.
    No random capitalization (i.e. no capitalizations in the middle of a word).
    No punctuation.
```

Pa ing Example

```
Family  Other Family
```

Failing Example

```
family  other family  other-family  FamilY  Family?
```

subfamily

Linnaean Subfamily. The formatting is:

```
    Capitalized first letter for each separate word.
    No random capitalization (i.e. no capitalizations in the middle of a word).
    No punctuation.
```

Pa ing Example

```
Subfamily  Sub Family
```

Failing Example

```
subfamily  sub family  sub-family  subFamilY  Subfamily?
```

tribe

Linnaean Tribe. The formatting is:

```
Capitalized first letter for each separate word.
No random capitalization (i.e. no capitalizations in the middle of a word).
No punctuation.
```

Pa ing Example

```
Tribe  Tribe Name
```

Failing Example

```
tribe  tribe name  tribe-name  tRibE  Tribe?
```

genus

Linnaean Genus. The formatting is:

```
Capitalized first letter for each separate word.
No random capitalization (i.e. no capitalizations in the middle of a word).
No punctuation.
```

Pa ing Example

```
Genus  Genus Name
```

Failing Example

```
genus  genus name  genus-name  gEnuS  Genus?
```

subgenus

Linnaean Subgenus. The formatting is:

```
Capitalized first letter for each separate word.
No random capitalization (i.e. no capitalizations in the middle of a word).
No punctuation.
```

Pa ing Example

```
Subgenus  Subgenus Name
```

Failing Example

subgenus subgenus name

Darwin Core: A brief phrase or a standard term ("cf.", "aff.") to express the determiner's doubts about the Identification.

recordedBy

The primary collector's name, should only be one name pair. The general format is:

```
Last,First
First Last
```

Either is accepted, however please make note of how the order changes when using a comma. The resulting value stored in the database will be of the format `Last,First`. Please do not enter a value if this is unknown, instead of iterations of `Unknown Unknown`

Pa ing Example

`Leopold,Aaron` `Aaron Leopold` `Leopold,Unknown` `Unknown Leopold`

Failing Example

`Leopold,aaron` `aaron leopold` `Aaron` `Unknown Unknown`

otherCollectors

The names of the collectors other than the primary collector, may be multiple names. The general format is:

```
Last,First
Last1,First1 | Last2,First2
Last1,First1|Last2,First2

First Last
First1 Last1 | First2 Last2
First1 Last1|First2 Last2
```

Please note: I am using numbers to indicate different names, however numbers are not allowed in this field.

The list separator must be the pipe character '|'. Either list format/naming order is accepted, however please make note of how the order changes when using a comma. The resulting value stored in the database will be of the format `Last,First|Last,First|Last,First`

Pa ing Example

`Leopold,Aaron  Aaron Leopold  Leopold,Aaron | Doe,Jane  Leopold,Aaron|Doe,Jane`

Failing Example

`Leopold  leopold,aaron  Leopold,aaron | doe,Jane  Leopold,Aaron, Doe,Jane`

identifiedBy

The name(s) of who identified the specimen, may be multiple names.

Darwin Core: A list (concatenated and separated) of names of people, groups, or organizations who assigned the Taxon to the subject.

The general format is:

```
Last,First
Last1,First1 | Last2,First2
Last1,First1|Last2,First2

First Last
First1 Last1 | First2 Last2
First1 Last1|First2 Last2
```

Either is accepted, however please make note of how the order changes when using a comma. The resulting value stored in the database will be of the format `Last,First` . Please do not enter a value if this is unknown, instead of iterations of `Unknown Unknown`

Pa ing Example

`Leopold,Aaron  Aaron Leopold  Leopold,Aaron | Doe,Jane  Leopold,Aaron|Doe,Jane`

Failing Example

```
Leopold  leopold,aaron  Leopold,aaron | doe,Jane  Leopold,Aaron, Doe,Jane
```

dateIdentified

The date in which the specimen was identified.

Darwin Core: The date on which the subject was determined as representing the Taxon.

This follows the general date format:

```
YYYY-MM-DD
YYYY-MM
YYYY
```

May not be future dates, or dates that preceed reasonable floor thresholds (e.g. less than 999)

verbatimDate

The verbatim date as it is written on the label (if present). This has no structured format as it is a record of whatever is hand written

General Example

```
Jan-Jun 2019
```

collectedYear

Year of collection. This follows the general date format, only concerning the year. I.e.: `YYYY`

May not be future dates, or dates that preceed reasonable floor thresholds (e.g. less than 999)

collectedMonth

Month (as a number, 1-12) of collection. This follows the general date format, only concerning the month. I.e.: `MM`

collectedDay

Day (as a number, 1-31) of collection. This follows the general date format, only concerning the day. I.e.: `DD`

dateEntered

The date in which the specimen was entered in the system. This follows the general date format:

```
YYYY-MM-DD
```

If left empty, an automated time stamp will be assigned on the insert query.

May not be future dates, or dates that preceed reasonable floor thresholds (e.g. less than 999).

sex

The sex of the specimen. Must match the ***first letter*** of one of the control values:

```
Male  Female  Gynandromorph
```

Darwin Core: The sex of the biological individual(s) represented in the Occurrence.

Pa ing Example

```
M  F  G
```

Failing Example

```
Male  Female  Gynandromorph
```

lifeStage

The life stage the specimen is in. Must match one (only one) of the control values exactly (case sensitive):

```
egg  larva  pupa  adult
```

Darwin Core: The age class or life stage of the biological individual(s) at the time the Occurrence was recorded.

habitat

Darwin Core: A category or description of the habitat in which the Event occurred.

There are no generalized formatting schemas for this field

General Example

```
oak savanna  pre-cordilleran steppe
```

occurrenceRemarks

Remarks from the collector. Not generalized formatting schemas for this field

Darwin Core: Comments or notes about the Occurrence.

molecularOccurrenceRemarks

Remarks from the collector. Not generalized formatting schemas for this field

samplingProtocol

The protocol in which the specimen was sampled. May be a list, where each item matches a control value exactly (case sensitive):

```
HandDirect  NetAerial  Light  LightUV  LightMV  LightMH  LightLED  LightOther  Bait
TrapMalaise  Trap
```

The format of the list follows previous field formatting, in that it must be delimited using the pipe '|' character.

Darwin Core: The name of, reference to, or description of the method or protocol used during an Event.

Pa ing Example

```
NetAerial  LightUV|LightLED
```

Failing Example

```
UV Light  LightUV,LightLED
```

```
example
```

country

Country of the collection. Must be in the list of accepted countries, matching exactly (case sensitive).

You may find the list here: `https://github.com/FLMNH-MGCL/spesql/blob/typescript-rewrite/src/renderer/assets/countries.ts` . This list is a TypeScript array of Country objects, containing a `name` and `code` field like so:

```
// generalized structure of a country object
type Country = {
  name: string;
  code: string;
};

// example country object
const exampleCountry = {
  name: 'Afghanistan',
  code: 'AF',
};
```

Please ensure the value matches the `name` field, **not** the `code` field

stateProvince

The State or Province of the collected specimen. Follows traditional proper noun formatting:

```
Capitalize first letter of each word.
No random capitalization.
```

county

The County of the collected specimen. Follows traditional proper noun formatting:

```
Capitalize first letter of each word.
No random capitalization.
```

municipality

The Municipality of the collected specimen. Follows traditional proper noun formatting:

```
Capitalize first letter of each word.
No random capitalization.
```

Darwin Core: The full, unabbreviated name of the next smaller administrative region than county (city, municipality, etc.) in which the Location occurs. Do not use this term for a nearby named place that does not contain the actual location.

Darwin Core Recommendation: Recommended best practice is to use a controlled vocabulary such as the Getty Thesaurus of Geographic Names.

General Example

```
Holzminden  Araçatuba  Ga-Segonyana
```

locality

The Locality data of the collected specimen. No generalized formatting in place.

Darwin Core: The specific description of the place. Less specific geographic information can be provided in other geographic terms. This term may contain information modified from the original to correct perceived errors or standardize the description.

General Example

```
Bariloche, 25 km NNE via Ruta Nacional 40 (=Ruta 237)
```

elevationInMeters

The elevation at which the specimen was located. Measurements will be converted to meters, however please ensure the unit is either:

`m` for meters, `mi` for miles, `ft` for feet

Pa ing Example

```
50m  1mi  100ft
```

Failing Example

```
50 meters  1 mile  1 foot  100
```

decimalLatitude

The latitude the specimen was located at. The value may be a floating point number (decimal) between -90 and 90.

```
-90 <= latitude <= 90
```

[redacted]

Failing Example

```
50 meters  1 mile  1 foot  100  0m  0
```

verbatimLatitude

The latitude as written on the label (if present). No generalized formatting, as this is just a record of what is hand written

verbatimLongitude

The longitude as written on the label (if present). No generalized formatting, as this is just a record of what is hand written

georeferencedBy

Darwin Core: A list (concatenated and separated) of names of people, groups, or organizations who determined the georeference (spatial representation) for the Location.

This follows the standard, list format:

```
Last,First
Last1,First1 | Last2,First2
Last1,First1|Last2,First2

First Last
First1 Last1 | First2 Last2
First1 Last1|First2 Last2
```

Pa ing Example

```
Brad Millen | Kristina Yamamoto | Janet Fang  Millen,Brad | Yamamoto,Kristina | Fang,Janet
```

Failing Example

disposition

The disposition (current state) of the specimen as it is currently in the collection. Must match one (only one) of the control values (case sensitive):

```
Present  Missing  Sample Used Up  On Loan
```

isLoaned

Indication of if the specimen is on loan. Must be one of the two accepted values (case sensitive):

`Y` for yes, `N` for no

loanInstitution

The name of institution specimen is loaned to. There is no generalized formatting for this field.

If the isLoaned field is indicated as `N` this field is required to be empty.

If the isLoaned field is indicated as `Y` this field is required.

loaneeName

Name of person responsible with specimen while loaned. This follows the general format:

```
First Last
Last,First
```

If the isLoaned field is indicated as `N` this field is required to be empty.

loanDate

Date specimen went on loan

If the isLoaned field is indicated as `N` this field is required to be empty.

If the isLoaned field is indicated as `Y` this field is required.

loanReturnDate

Expected return date for specimen, update to actual when returned.

This follows the general date format:

```
YYYY-MM-DD
YYYY-MM
YYYY
```

May not preceed reasonable floor thresholds (e.g. less than 1990)

preparations

The preparations of the specimen. Must match one of the control values (case sensitive)

`Wing Voucher` `Molecular Collection` `Pinned Collection` `Larval Collection` `Genetic Collection`

Darwin Core: A preparation or preservation method for a specimen.

freezer

The freezer the specimen is stored in (if applicable). The general format is as follows:

```
'Kawahara' followed by two digits:
Kawahara##
```

Pa ing Example

```
Kawahara05  Kawahara11
```

Failing Example

```
kawahara 5  karahara30  Kawahara119
```

rack

The rack on/in which the specimen is stored. The general format is:

```
1–3 characters long.
No punctuation.
Alphanumeric
```

Pa ing Example

```
B32
```

box

Box number of specimen. The general format is:

```
Integer value between 1–99.
1 <= box <= 99
```

tubeSize

The size of the tube/container/other used. It must match one (only one) of the control values:

```
papered  50falcon  15falcon  microcentrifuge
```

associatedSequences

A list of identifiers (publication, bibliographic reference, global unique identifier, URI) of literature associated with the Occurrence.

Note: This field is too complex to validate programmatically, so please refer to the example below to see the manner in which you should structure the list.

General Example

```
 http://www.sciencemag.org/cgi/content/abstract/322/5899/261, Christopher J. Conroy, Jennifer L. Neuwald. 2008. Phylogeographic study of the California vole, Microtus californicus Journal of Mammalogy, 89(3):755–767., Steven R. Hoofer and Ronald A. Van Den Bussche. 2001. Phylogenetic Relationships of Plecotine Bats and Allies Based on Mitochondrial Ribosomal Sequences. Journal of Mammalogy 82(1):131–137. | Walker, Faith M., Jeffrey T. Foster, Kevin P. Drees, Carol L. Chambers. 2014. Spotted bat (Euderma maculatum) microsatellite discovery using illumina sequencing. Conservation Genetics Resources.
```

associatedReferences

A list of identifiers (publication, global unique identifier, URI) of genetic sequence information associated with the Occurrence.

The general format follows the standard list format, and is as follows:

```
  link
  linkOne|linkTwo
  linkOne | linkTwo
```

Pa ing Example

```
 http://www.ncbi.nlm.nih.gov/nuccore/U34853.1 |
http://www.ncbi.nlm.nih.gov/nuccore/GU328060 |
http://www.ncbi.nlm.nih.gov/nuccore/AF326093
```

Failing Example

```
 http://www.ncbi.nlm.nih.gov/nuccore/U34853.1,
http://www.ncbi.nlm.nih.gov/nuccore/GU328060,
```

`http://www.ncbi.nlm.nih.gov/nuccore/AF326093`

withholdData

Indication of if there was data withheld. Must be one of the two accepted values (case sensitive):

`Y` for yes, `N` for no

Note: please refer to informationWithheld in Darwin Core to view more about what this field is

reared

Indication of if the specimen has been reared. Must be one of the two accepted values (case sensitive):

`Y` for yes, `N` for no

recordEnteredBy

The (spesql) credentials of the individual inserting this into the database. This is automatically determined, **DO NOT** manually attempt to insert this field. The insertion will fail.

This field is available for logging and querying.

modifiedInfo

The edit history of this entry in the database. This is automatically determined and updated as changes to the entry are made, **DO NOT** manually attempt to insert this field. The insertion will fail.

This field is available for logging and querying.

fieldNotes

Notes from field, other notes about specimen. There is no specified formatting to follow