

Rapid Data Entry (RDE) in EMu

Last modified: 20 Sep 2023

Overview

Rapid Data Entry in EMu provides users with a one-tab option to enter minimal collection data about a specimen without using the Collection Events and Sites modules. A second tab has also been added, called OCR Text, which contains OCR text of the specimen label that can be used to query for similar records. This guide assumes you have basic EMu training already.

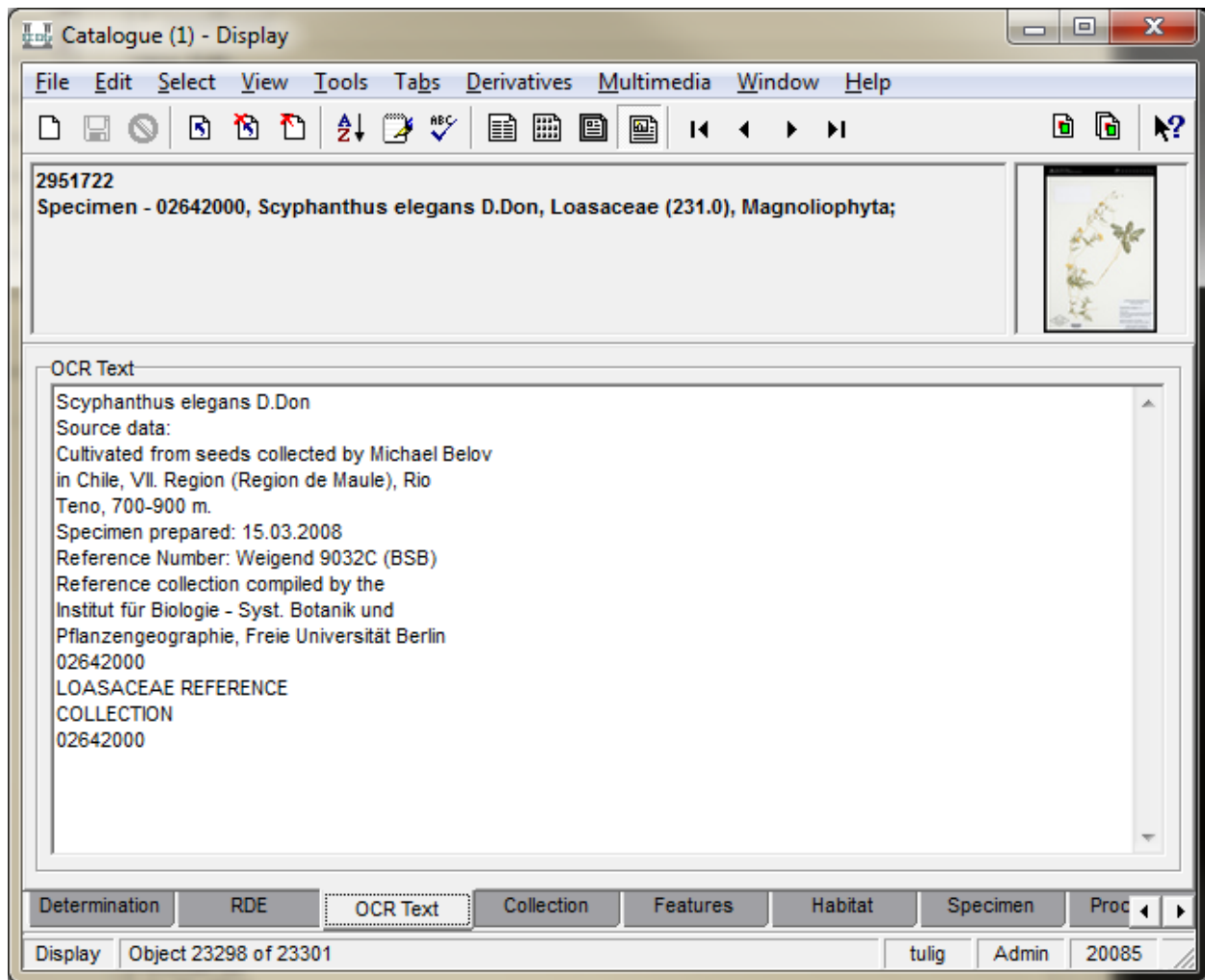
RDE Tab Example

The screenshot shows the EMu software interface with the 'RDE' tab selected. The window title is 'Catalogue (1) - Display'. The menu bar includes File, Edit, Select, View, Tools, Tabs, Derivatives, Multimedia, Window, and Help. The toolbar contains various icons for file operations and navigation. The main content area displays the specimen ID '2951722' and the specimen name 'Specimen - 02642000, *Scyphanthus elegans* D.Don, Loasaceae (231.0), Magnoliophyta;'. A small image of the specimen is visible on the right. Below the main content, there are several input fields for data entry:

- RDE Status:** Processing Status: Unprocessed (dropdown menu)
- RDE Collection Events Details:** Collector, Collection Number Prefix, Collection Number, Collection Number Suffix, Collection Team, Date From, Date To, Verbatim Date.
- RDE Sites Details:** Country, Province/State.
- RDE Sites Details (continued):** County, City, Precise Location, Township, Range, Section, Latitude (Dec.), Latitude (DMS), Longitude (Dec.), Longitude (DMS), Elevation From (m/ft), Elevation To (m/ft).

The bottom of the interface shows a tabbed navigation bar with 'RDE' selected, and other tabs for Determination, OCR Text, Collection, Features, Habitat, Specimen, and Proc. The status bar at the bottom indicates 'Display Object 23298 of 23301' and shows user information 'tulig Admin 20085'.

OCR Text Tab Example



Getting Started

Search for Records to Database

Query Defaults

You need to set query defaults to search for records that are not yet databased. This includes records without an attachment (\\!\+) in the RDE Collector field or the Collection Event field. RDE default values have been set up for select taxonomic groups and/or digitization projects and can be loaded by going to:

- Edit->Default Values->Change->RDE default values – Algae
- Edit->Default Values->Change->RDE default values – Biocultural
- Edit->Default Values->Change->RDE default values – Bryophytes
- Edit->Default Values->Change->RDE default values – Endless Forms
- Edit->Default Values->Change->RDE default values – Ferns

- Edit->Default Values->Change->RDE default values - GLOBAL
- Edit->Default Values->Change->RDE default values - Lichens
- Edit->Default Values->Change->RDE default values – Macrofungi
- Edit->Default Values->Change->RDE default values – Microfungi
- Edit->Default Values->Change->RDE default values – Vascular plants

***There is a difference between Query Defaults and Insert Defaults. You can only load default values for the mode you are in. Make sure you are in Query mode to load these default values.**

***REMEMBER to change back to “No Defaults” when working on another project**

- Edit->Default Values->Change->No Defaults

Use OCR Text to Find Similar Records

The OCR text can be used:

- To search for records specific to your project, or
- To group records with similar labels so ditto can be used between records, or
- A combination of the two.

Examples:

For GLOBAL, staff must transcribe label data for all lichen and bryophyte specimens from all geographic filing regions except “US & Canada.” Records could be grouped by a specific collector and/or location.

For All Asia, staff must transcribe full label data for vascular plants collected from specific countries within the geographic filing regions Tropical Asia, Northern Asia, and Pacifica.

For Africa, staff must transcribe full label data for vascular plants collected from specific countries within the geographic filing region Africa.

Any full words can be used on the same line in query mode to find those words anywhere in the OCR text for any given record or set of records. When searching for terms in a string, use quotes around the string. **Remember, you must escape the quotes (or any punctuation in a query) with a backslash.**

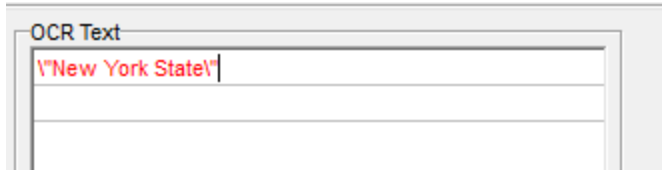
Examples:

`\ "FLORA OF THE PHILIPPINES\ "`

`\ "PLANTS OF CAMEROON\ "`

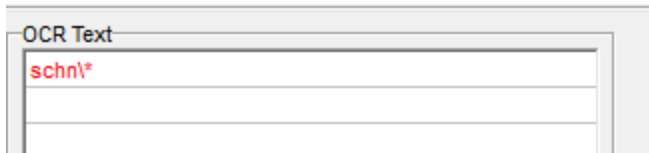
`\ "long island\ "`

`\ "COSTA RICA\ "`

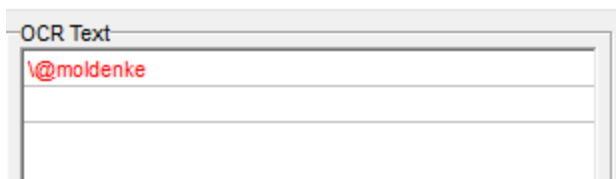


If you are not finding good results using full words, try partial matches. Check the OCR text for common misspellings and query by the misspelling. You can also use partial matches by starting or ending a phrase with *.

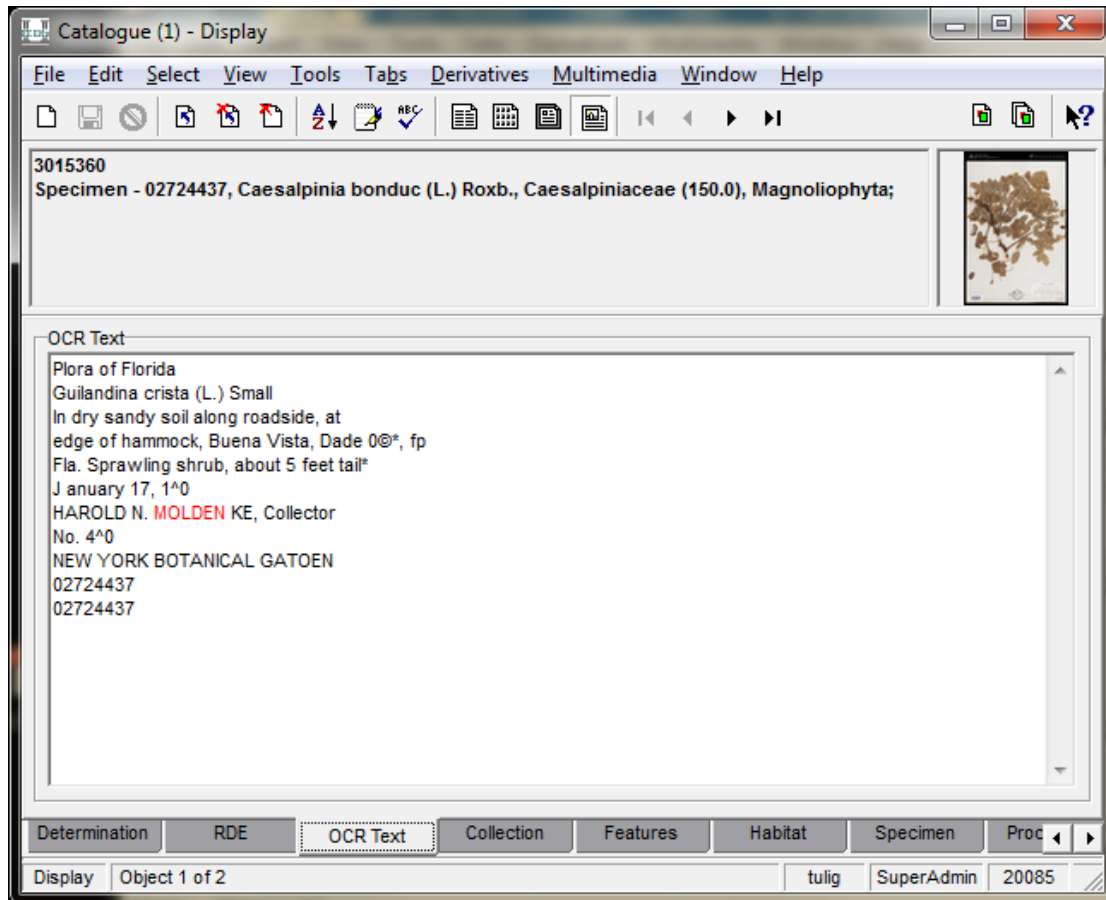
Try `schn*` when looking for “Schneider” (but you will also find Schnurmann, Schnorrenberger, etc.)



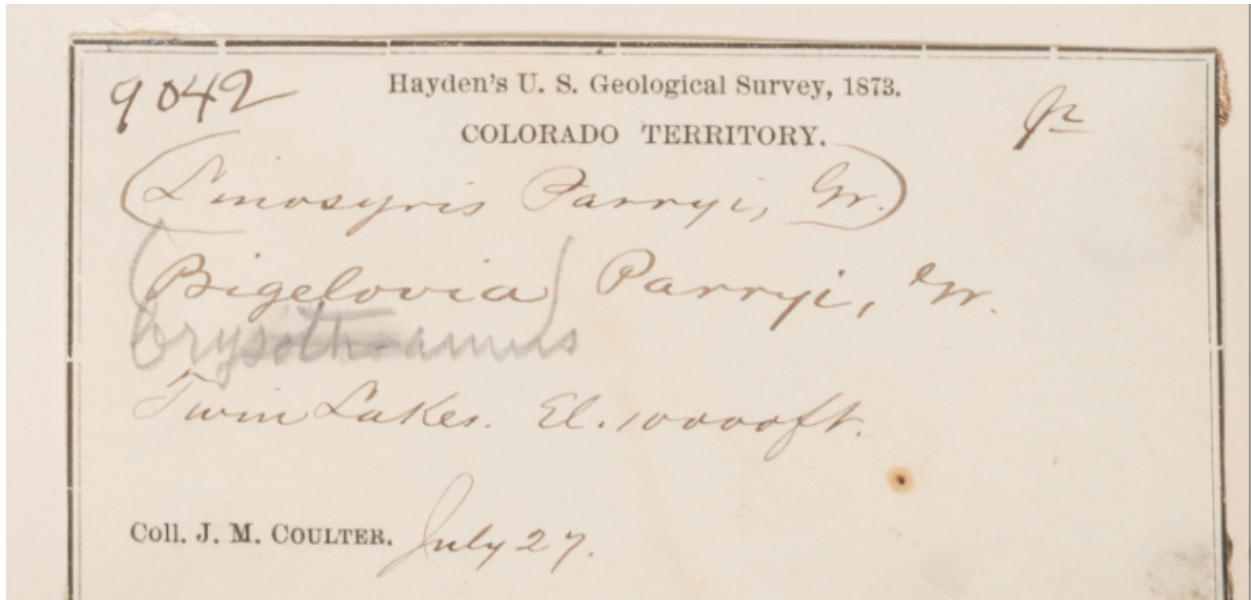
Try sounds like matches (\\@) to catch variations of useful words that the OCR got wrong.



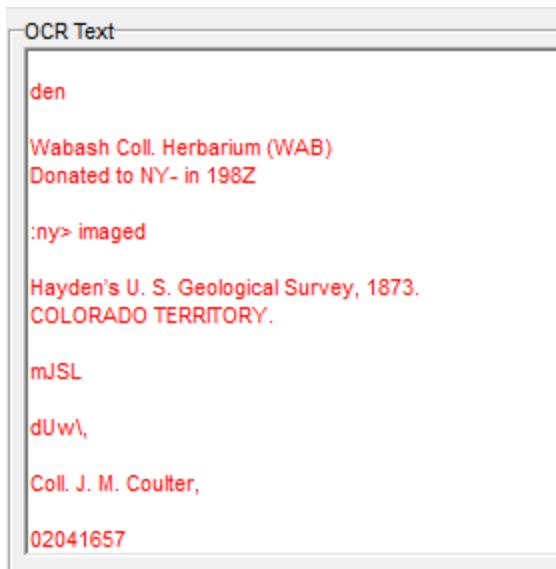
`\\@moldenke` finds potential Moldenke records with OCR translation errors.



For handwritten labels, the header and/or footer of the label is often typewritten and can be used to query for certain projects, regions, or collectors.



In this example, the OCR of the header and footer can be used to find the survey name, the date, the state and the collector name even though most of the label is handwritten.



If a query finds too many matches, look for a term in the OCR text that may reduce the records further and run another query with the values separated by spaces all on the same line:



***Keep the values all on the same line to search OCR for ALL terms to run an AND search.**

***Putting values on separate lines runs an OR search and will produce results that may have only one of the terms.**

This limited results from Colorado, collected by Goodrich all with Maybell in the precise location. Keep in mind that the more terms you add, the more likely one of the words has an OCR translation error in it, so running queries with all variations is helpful.

RDE Records with Countries and States Pre-Populated

Some records may have the Country and/or State fields already populated with values. These came from two sources:

1. OCR text matches for location values

We have tried to automate the population of these fields based on phrases in the OCR text, like “Plants of Colorado.” **In some cases, the values in these fields could be incorrect. Please update these values when you are checking the images and do not assume you should keep these values in the record.**

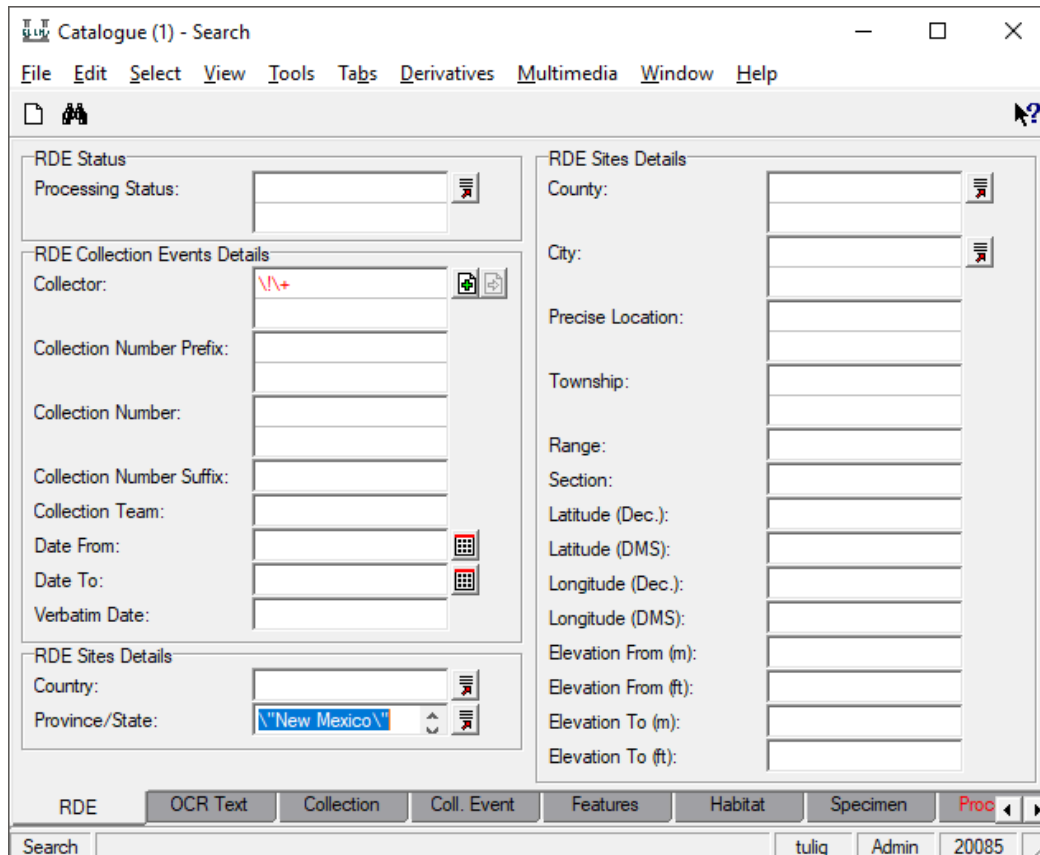
2. Geography Crowdsourcing

For records that are mostly handwritten or difficult to use OCR to determine localities, we have been using crowdsourcing to add countries and states to RDE records. You can also use the RDE state and country fields to query for records from the same regions to increase efficiency. You can query this field on its own or use this in conjunction with terms on the OCR tab.

How to Find Records with Country and State values Pre-Populated

To find specimens from New Mexico that aren’t transcribed yet, just add the state on the RDE tab in query mode and you’ll get all the records where New Mexico has already been added. Then just database the rest of the fields to complete the record.

Please remember, states and countries have only been added for a fraction of the records. Please check the OCR tab on these records and look for any terms that might help you find similar records.



Database Records using the RDE Tab

Once you find a group of similar records to transcribe, enter all the information on the label that corresponds to an RDE field. If the field does not exist on the RDE tab, but you must enter it for your project, YOU CANNOT USE THE RDE tab. You must database a record using the Collection Events and Sites modules. Most Herbarium digitization grants can limit their data entry to RDE only and skip any missing fields, but please check with your supervisor first.

RDE List View

Summary Data does not contain RDE fields, so it can be difficult to quickly see in a list which records have been completed. However, you can easily view the fields by using different views.

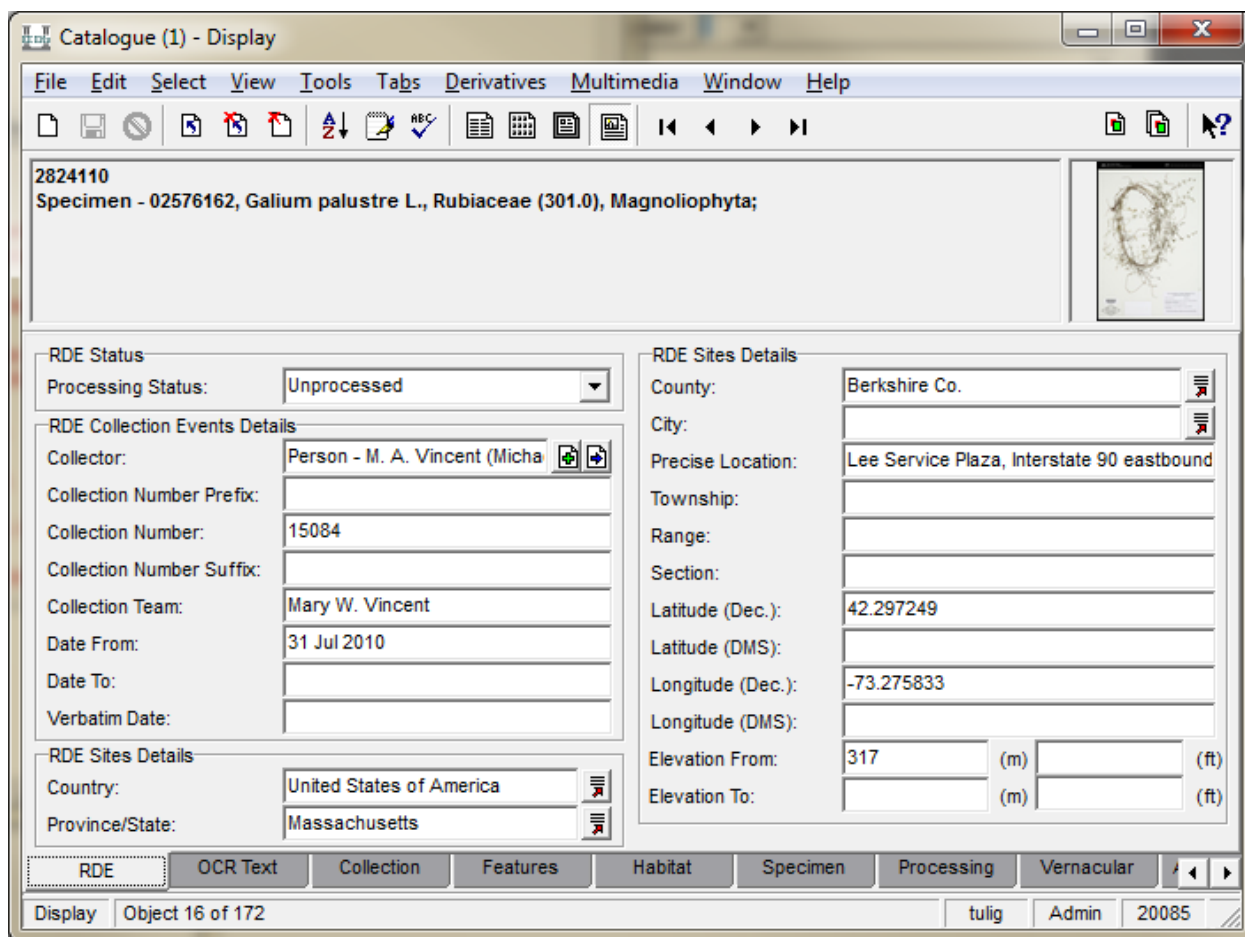
View->List Settings->Choose List->RDE fields

Collection E...	Barcod.	Family: (Versio...	Scientific Name: (Ve...	Collector: (RDE Collection Events Details)	Collection	Date From: (RDE C...	Country: (RDE Sites Details)	Province/State: (RDI
1	02576250	Cistaceae	Lechea mucronata ...	Person - M. A. Vincent (Michael A. Vincent); ...	15098	31 Jul 2010	United States of America	Massachusetts
2	02578363	Lamiaceae	Thymus serpyllum L.	Person - M. A. Vincent (Michael A. Vincent); ...	15151	06 Aug 2010	United States of America	Pennsylvania
3	02576136	Rosaceae	Rosa palustris Mars...	Person - M. A. Vincent (Michael A. Vincent); ...	15082	31 Jul 2010	United States of America	New York
4	02576126	Poaceae	Panicum virgatum L.	Person - M. A. Vincent (Michael A. Vincent); ...	15078	31 Jul 2010	United States of America	New York
5	2069601	Asteraceae	Crepis tectorum L.	Person - N. H. Holmgren (Noel H. Holmgren); ...	9774	20 Oct 1980	United States of America	New York
6	2443538	Caprifoliaceae	Sambucus pubens ...	Person - E. W. Wood (Emily W. Wood); Deter...	4217	18 Jun 1979	United States of America	Pennsylvania
7	1895862	Asteraceae	Anthemis cotula L.	Person - F. H. Utech (Fredrick H. Utech); Auth...	323	30 Jun 1979	United States of America	Pennsylvania
8	02576148	Rosaceae	Spiraea alba Du Roi	Person - M. A. Vincent (Michael A. Vincent); ...	15088	31 Jul 2010	United States of America	Massachusetts
9	02576175	Rubiaceae	Galium glaucum L.	Person - M. A. Vincent (Michael A. Vincent); ...	15072	30 Jul 2010	United States of America	New York
10	02576147	Asteraceae	Eupatorium perfoliat...	Person - M. A. Vincent (Michael A. Vincent); ...	15117	05 Aug 2010	United States of America	Connecticut
11	1465366	Fagaceae	Quercus rubra L.	Person - M. C. Pace (Matthew Charles Pace);...	485	30 Jul 2009	United States of America	New York
12	02576143	Rosaceae	Rosa rugosa Thunb.	Person - M. A. Vincent (Michael A. Vincent); ...	15107	01 Aug 2010	United States of America	Rhode Island
13	02576132	Asteraceae	Tanacetum vulgare L.	Person - M. A. Vincent (Michael A. Vincent); ...	15104	01 Aug 2010	United States of America	Rhode Island
14	1884540	Asteraceae	Ambrosia trifida L.	Person - F. H. Utech (Fredrick H. Utech); Auth...	370	10 Aug 1979	United States of America	Pennsylvania
15	1925496	Asteraceae	Bidens cernua L.	Person - N. H. Holmgren (Noel H. Holmgren); ...	9767	27 Sep 1980	United States of America	New York
16	02576162	Rubiaceae	Galium palustre L.	Person - M. A. Vincent (Michael A. Vincent); ...	15084	31 Jul 2010	United States of America	Massachusetts
17	2093223	Asteraceae	Erigeron philadelphi...	Person - N. H. Holmgren (Noel H. Holmgren); ...	9838	31 May 1981	United States of America	New York
18	1465367	Fagaceae	Quercus rubra L.	Person - M. C. Pace (Matthew Charles Pace);...	120	27 Apr 2008	United States of America	New York
19	2546611	Ericaceae	Vaccinium macrocar...	Person - B. E. Hall (Bruce E. Hall)	2007	18 Jul 2007	United States of America	Maine
20	1465369	Fagaceae	Quercus rubra L.	Person - M. C. Pace (Matthew Charles Pace);...	482	16 May 2009	United States of America	New York
21	02576168	Brassicaceae	Raphanus raphanist...	Person - M. A. Vincent (Michael A. Vincent); ...	15068	30 Jul 2010	United States of America	New York

Display Mode

View Thumbnails

To open an image without going to the Multimedia tab, turn on View Thumbnails by going to View->Thumbnails. You can double click on the image from any tab to open the file in your default image viewer.



RDE Tab Data Entry

Using the fields on the RDE tab, enter all information from the label.

Processing Status

This field is populated by Admin users and can be ignored by project staff.

Collector

Mandatory field: must be populated

For US collections, most collectors should already exist in the Parties module. For collections from Asia and Africa, this may not be true. It will take some time to get used to how names appear on labels, but always try to select an existing record if you feel there could be a match rather than create a new one. You may find it more efficient to keep the Parties Module open to query the various name fields.

If no collector is specified or is illegible use one of the following Parties records:

- Collector unspecified (IRN 20106) = collector is not indicated on the label
- Illegible collector name (IRN 20104) = collector indicated, but completely illegible.
- see Collection Note (IRN 103805) = collector indicated but is partially legible or it is not possible to determine which Parties record to assign (e.g. collector initials only). Record verbatim within square brackets [] the text you can decipher in the “Collection Notes” field, including within quotation marks the text transcribed verbatim from the label.

If you need to add a new record to the Parties module, please follow the [Quick Guide to Entering a New Parties record](#).

Collection Team

Enter any remaining collectors as a string in the same order on the label with the participants’ first initials only. Use commas to separate each name. No need to add “and” or “with” in the string.

Example: X. Aguirre, R. Miranda, C. Robles

Collection Number Prefix, Number, Suffix

The collection number is usually a numeric value assigned by the primary collector: S. A. Mori 1145. The Collection Number field can only accept numeric values. Occasionally, the collection number will include additional values before or after the number which should be entered in the Collection Number Prefix or Collection Number Suffix fields. Any punctuation surrounding the collection number should also be entered in the prefix or suffix fields.

If there is no collection number on the label, please enter “s.n.” (no spaces and without quotes) in the Collection Number Suffix field.

Example: 1145a

Collection Number: 1145

Collection Number Suffix: a

Example: 1995-12345

Collection Number Prefix: 1995-

Collection Number: 12345

Collection Dates: Date From, Date To, Verbatim Date

The fastest way to enter a collection date is by using the keypad to enter the date in all digits in the format of DAY MONTH YEAR.

Almost all of our collections will be from the 1800s-2000s. If you see dates outside this range on a label, they are probably wrong and should be checked by a project manager. Specimens from the 1700s are possible, but VERY rare, so please double-check before entering any dates from the 1700s.

****YOU MUST ENTER THE FULL YEAR IN ALL DATE FIELDS. IF LABEL JUST HAS '97 FOR THE YEAR, YOU MUST ENTER 1997 (OR 1897, ETC.)****

Date From

Most labels will only have one date and it should always be entered in the Date From field. The field forces data into a single format of DAY MONTH YEAR. You can enter in various formats and it will automatically update it to the correct format. The Date From field can also accept dates as just a year or a month and a year. Please put any part of the date that can be standardized into the Date From field and the rest in Verbatim Date.

The Verbatim Date is just a text field making it difficult to effectively use dates in searches. If the label says "6/18/99" and this is all that is entered in Verbatim Date, we won't easily be able to find it when searching for a year. It is usually easy to guess if the correct year is 1899 or 1999. Only in extraordinarily rare cases would we have a specimen from 1799. You can still enter "6/18/99" in the Verbatim Date field, but in Date From, please enter "18 Jun 1899" if you are fairly certain the date is 1899.

Interns should ask for help to tell the difference between 1800's and 1900's labels until they get the hang of it. Please remember most collectors have a birth date in their records and you can use that as an indication of what the likely collection date is.

Examples:

Information on label: early May 1884

Enter in EMu:

Date From: May 1884

Verbatim Date: early May 1884

Information on label: spring 1884

Enter in EMu:

Date From: 1884

Verbatim Date: spring 1884

Roman numerals in dates are typically used to indicate the month collected.

Enter in EMu:

Information on label: 5 VII 75

Enter in Date From: 5 Jul 1975


No matter how the date is written or typed on a label, if you are fairly certain what the date is, there is no need to enter the formatting of the date as on the label in the Verbatim Date field. Especially for things like “4 August ‘97” just enter “4 Aug 1897” in the Date From field and nothing in the Verbatim Date field.

Please enter “s.d.” (no spaces and without quotes) in the Verbatim Date field when no date is indicated at all. There is nothing to put in Date From for this.

The Verbatim Date field should only be used to record information that is not an actual date (“spring”) or when the format of the date used on the label when you truly cannot distinguish between the day and the month.

Geographic Hierarchy: Country, Province/State, County, City

Each field in the geographic hierarchy is part of a lookup list with values stored in EMu for most regions.

All lookup lists have this icon  next to it. You can click on the icon to select a value from the list, or you can start entering some text and click F12 to auto-complete the field.

Please choose existing values from the lookup list, with the possible exception of “County” (see that section below). Only use “Unknown” for Country. Leave any other value in the geographic hierarchy blank to indicate an unknown.

Country

Values for Country represent present-day names that have been reviewed and locked by the GIS Lab Manager. Please select a value from the Lookup List that corresponds to the country indicated on the collection label. If the collection label indicates a country that is no longer recognized, use the current country name instead. If the country is not specified, choose “Unknown” for the country.

Example: British Honduras is now known as Belize. Please enter Belize into the country field.

See here a [list of historical country names related to present-day names](#).

Province/State

Values for Province/State represent present-day names used for the next smaller administrative region below that of country, and that have been reviewed and locked by the GIS Lab Manager. Please select a value from the Lookup List that corresponds to the Province/State indicated on the collection label.

If the collection label indicates a Province/State that is not available in the list, verify the names and spelling on the collection label are current in Statoids (<http://www.statoids.com/>). If the name is no longer recognized, and the new name maintains the same political boundaries as the old name, then select the new name from the list. If you are not sure, or if the boundaries have changed, leave the Province/State field blank, and transcribe verbatim the Province/State name in the field “Precise Location.”

Leave the field blank if a state/province is not specified on the label.

If you have any questions, please ask a manager for help.

County

Values for County represent present-day names used for the next smaller administrative region below that of Province/State. These are currently being reviewed and standardized by the GIS Lab Manager.

Please select a value from the Lookup List that corresponds to the county/municipio indicated on the collection label.

If the county does not appear on the label, but you have information that clearly indicates the correct county, please enter a county. **Otherwise, leave it blank.**

Please consult this growing [list of Countries](#) for which the County values have been cleaned and standardized and for which you should select existing values only (i.e. do not add any new values):

All other countries are in the process of being cleaned and standardized and for which you may add values, if needed.

City

Values for City have not been thoroughly reviewed and cleaned. Please populate this field with caution and according to these instructions:

- Select or add to the Lookup List for the City field a named city or town if and only if the city or town is indicated on the collection label **AND** the specimen was collected within the bounds of that city or town (i.e. not near, in the vicinity of, on the outskirts of, etc.)
- Select or add to the Lookup List for the City field if and only if the Lookup List values for County, Province/State, and Country have been populated first.

If you are not certain the named place is an official city or town, enter the names in the Precise Location.

Keep in mind that most cities and towns have greatly changed in size over time, whereas counties mostly have not. **It is best not to add a name to the lookup list unless you are positive.**

Precise Location

Enter verbatim any text string that indicates where this specimen was collected, in the original language on the label and including appropriate diacritics and symbols where indicated. Do not translate location data. See [how to enter symbols and diacritics](#) and more information below under heading “Symbols and Diacritics.”

If a Province/State, County, or City was indicated on the collection label but not captured using the appropriate Lookup List(s), transcribe here the name(s) verbatim as they appear on the label, in hierarchical order, prior to transcribing the more precise location data.

Please try not to enter habitat information here like forest, prairie, dunes, etc. UNLESS it is related to a named place like, "Washington National Forest" or "Shepard's Hills," etc. However, if habitat is part of the locality string and it is too hard to parse from the rest of the locality, it is fine to enter it in the precise location.

If you copy text from the OCR tab, be sure to edit the text and delete any hidden characters (tab, carriage return, new line, line feeds).

Township, Range, Section

These fields apply only to collections from the United States. See [Public Land Survey System](#) for more details.

Township is a 36 square mile plot of land broken into 36 one square mile sections. The Range is the east or west distance from the principal meridian.

Each value should be entered in EMu as:

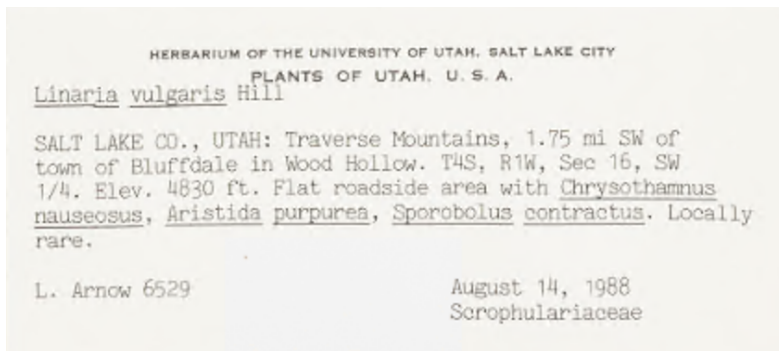
T<value><direction>

R<value><direction>

S<value>

Sometimes the section (S) can be difficult or have other associated values, so it might be difficult to strictly follow. Can be written as S11 NE1/4 or S1, S2, S3 (if 3 sections given) or what have you. If no section is included, leave it blank.

Example:



Enter in EMu as:

Township: T4S

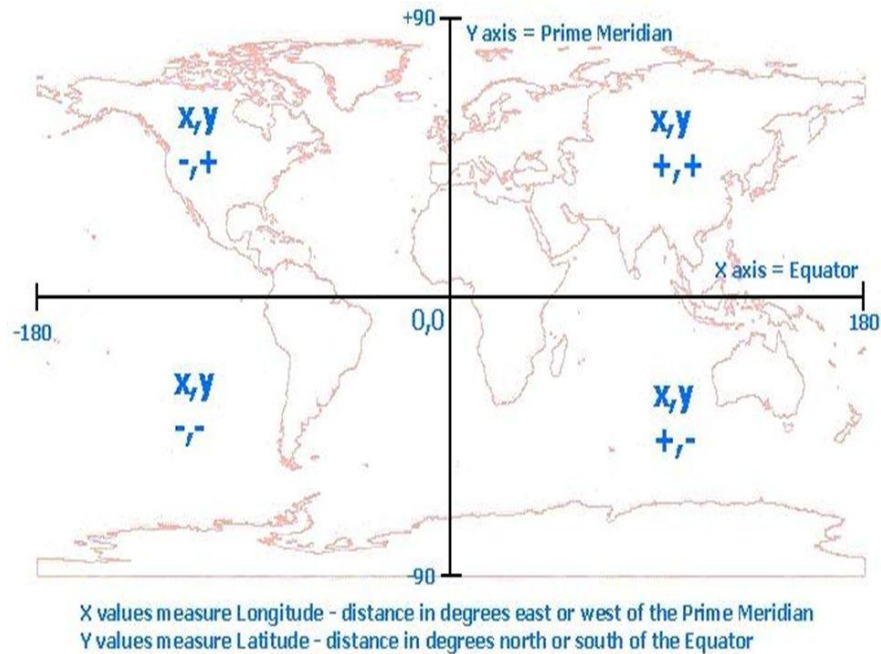
Range: R1W

Section: S16 SW1/4

Latitude and Longitude

Latitude/Longitude (Dec.): use if coordinates appear on the label as decimal degrees. Decimal degrees must be entered as a positive or negative value, NOT with a direction (N, S, E, W)

Positive latitudes are north of the equator, negative latitudes are south of the equator. Positive longitudes are east of Prime Meridian, negative longitudes are west of the Prime Meridian.



Latitude/Longitude (DMS): use if coordinates appear on the label as degrees, minutes, seconds. This includes if values are a combination of degrees, minutes and decimal minutes or decimal seconds. Values must include a direction (N, S, E, W).

How to enter variations in latitude and longitude in EMu

Example 1

N 44.49567° W 103.836933°

Label has decimal degrees, but the direction is indicated by N/W not by positive/negative

How to Enter in EMu:

Latitude (Dec.):	44.49567
Latitude (DMS):	
Longitude (Dec.):	-103.836933
Longitude (DMS):	

Example 2

N 44° 29' 46.65" W 103° 50' 15.44"

DMS with decimal seconds, direction entered BEFORE the the coordinates

How to enter in EMu:

Latitude (Dec.):	
Latitude (DMS):	44 29 46.65 N
Longitude (Dec.):	
Longitude (DMS):	103 50 15.44 W

Example 3

N 38 deg. 18.333 min.; W 113 deg. 54.384 min.

DMS with decimal minutes, direction entered BEFORE the the coordinates

How to enter in EMu:

Latitude (Dec.):	
Latitude (DMS):	38 18.333 N
Longitude (Dec.):	
Longitude (DMS):	113 54.384 W

Example 4

Ca 7.9 mi W of Knowles, 35 15'11", 105 02'32"

This is from a label specified as from Oklahoma. Lat/long is given in DMS, but no direction (N, S, E or W) is indicated. However, since we know the precise location, you can determine the correct direction:

Latitude (Dec.):	
Latitude (DMS):	35 15 11 N
Longitude (Dec.):	
Longitude (DMS):	105 2 32 W

Coordinate Ranges & UTM Values

If you come across a range of coordinates or UTM values, please enter them in the Precise Location field AFTER the location string. Please add "Coordinate Range:" before the range and leave the coordinate fields of the RDE tab blank. For UTM Values, please add "UTM:" before the values. These will get parsed into the correct fields after Sites records are created.

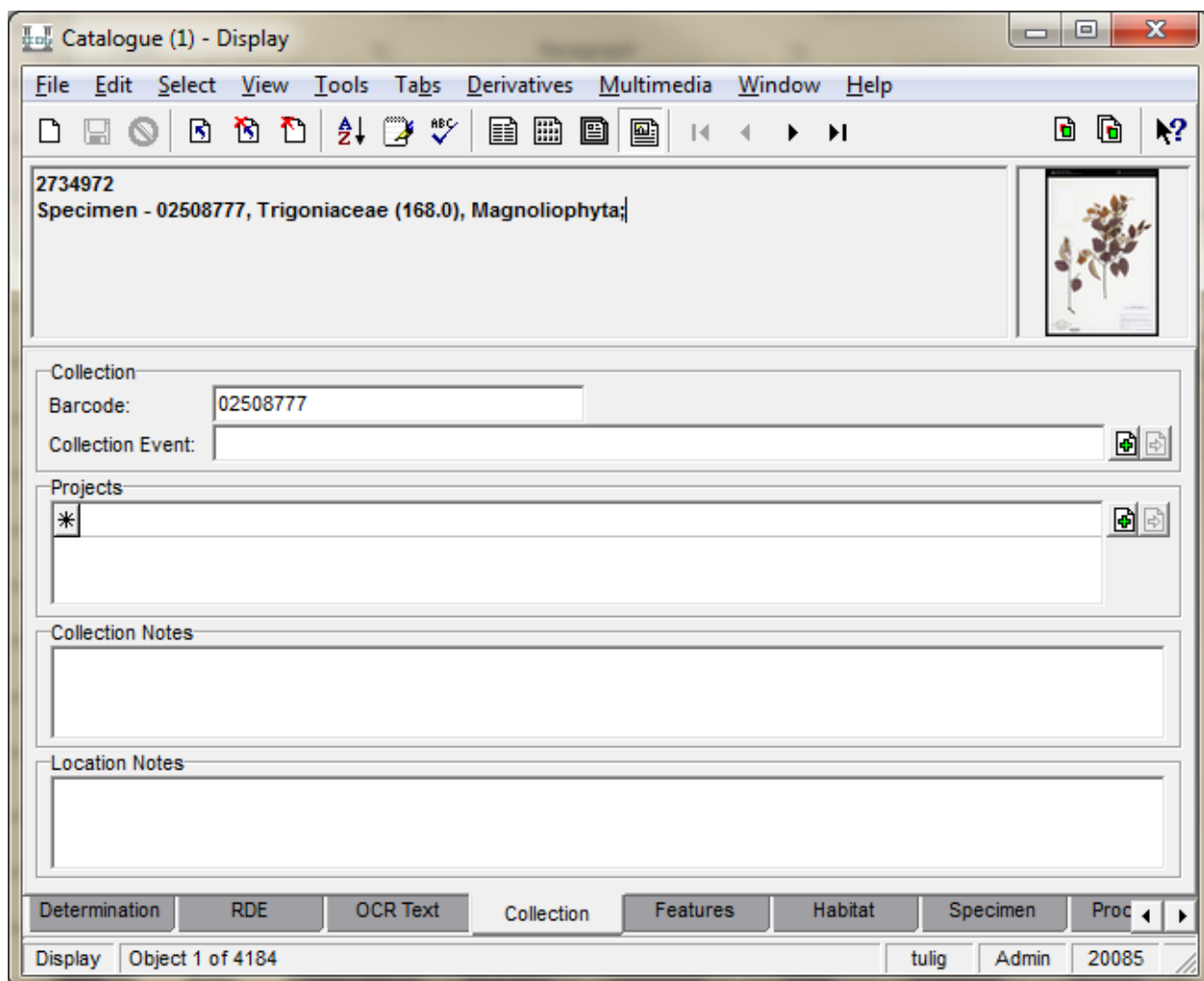
Elevation

Elevation should be entered as it appears on the label either in meters or feet in the corresponding fields. If an elevation range is not provided, leave the field "Elevation To" blank.

Notes

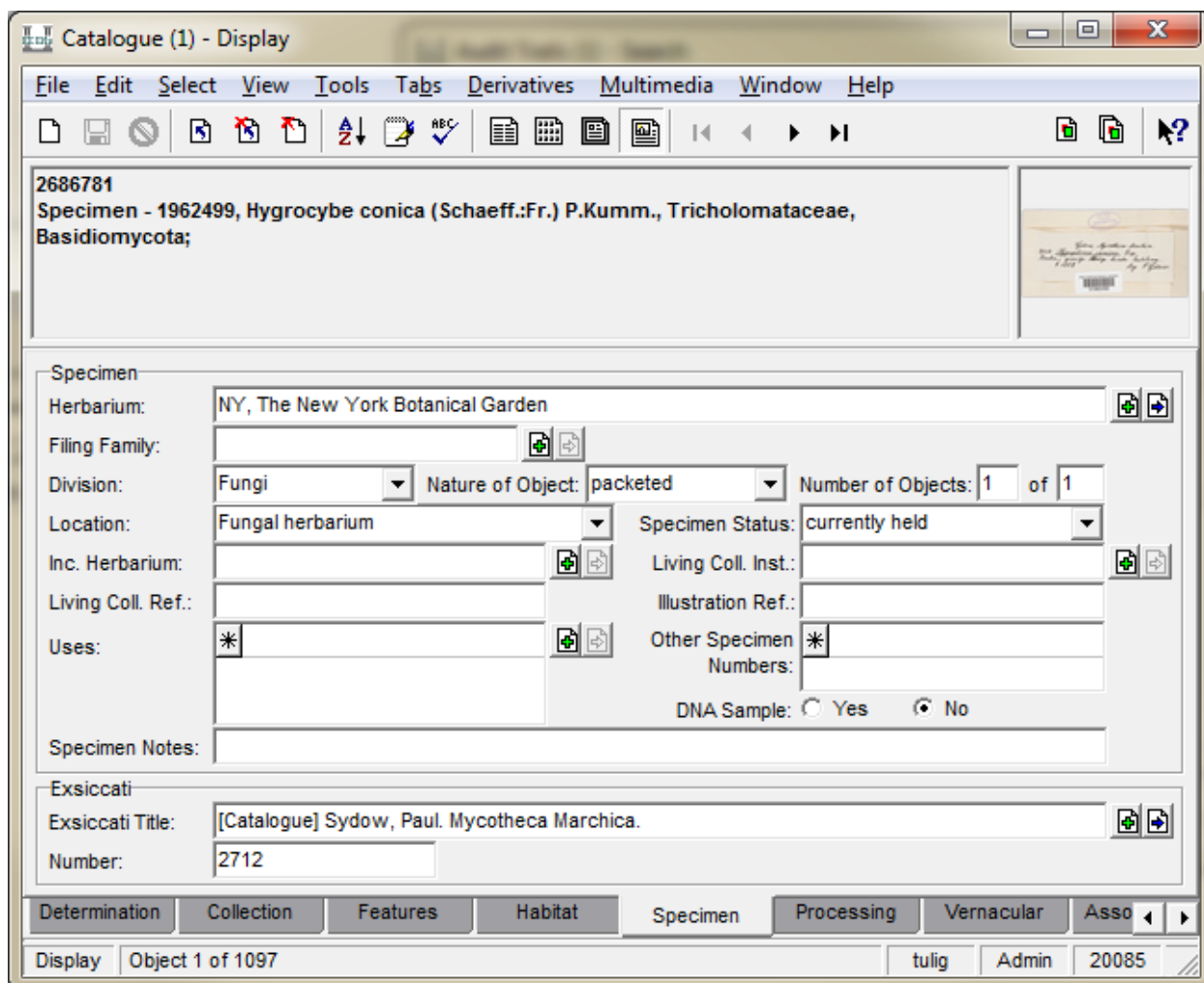
If you are uncertain about any collection or locality information, use the Collection tab to enter that information. Use either the Collection Notes or Location Notes fields, or both.

Type your notes in square brackets []. Enclose any text transcribed verbatim in quotation marks.



Exsiccatae

The Specimen tab is still available for staff needing to enter exsiccatae. Please ask your supervisor for instructions on how to do this correctly.



Features and Habitat Tabs

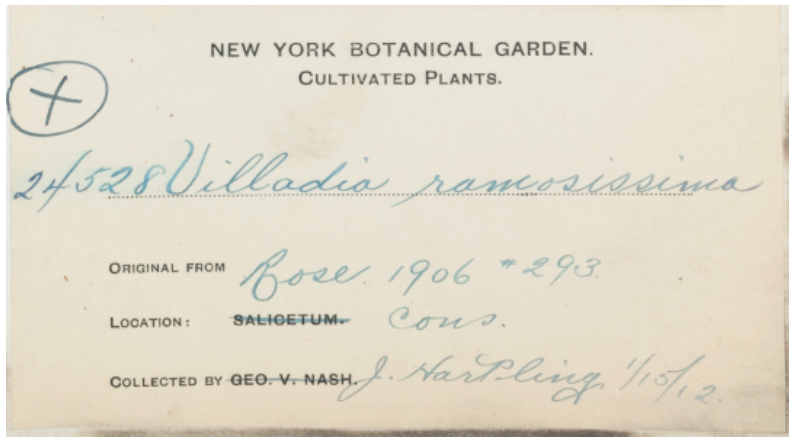
Most digitization grants only need to fill out the RDE tab. For other projects, you can still use the RDE tab in conjunction with any other tab in the Catalogue module, like the Features and Habitat tabs. Please check with your supervisor for instructions.

Cultivated Plants

When a specimen was collected from a plant grown in cultivation, e.g., in a botanical garden, arboretum, someone's yard, then:

- Capture on the RDE tab, the name of the collector, collection number, collection date, and location associated with the plant where it was collected in cultivation
- Capture verbatim and in quotation marks in "Collection Notes" the data related to the origin of the plant prior to cultivation.
- Select "Yes" in the field "Cultivated" on the Features tab

Example



Catalogue (2) - Edit

File Edit Select View Tools Tabs Derivatives Multimedia Window Help

3911866
Specimen - 03674884, *Villadia ramosissima* Rose, Crassulaceae (134.0), Magnoliophyta:

RDE Status
Processing Status: Unprocessed

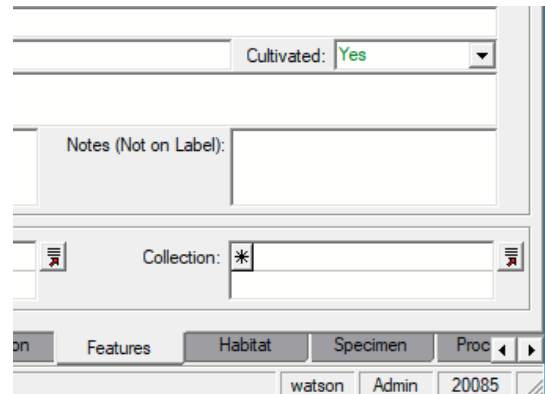
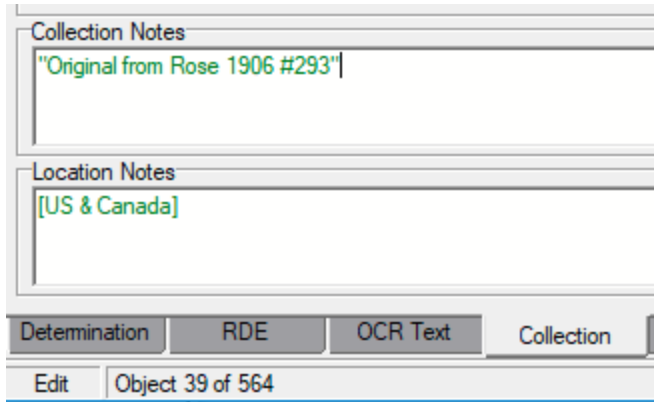
RDE Collection Events Details
Collector: Person - J. Hartling (John Hartling): Collec
Collection Number Prefix:
Collection Number: 24528
Collection Number Suffix:
Collection Team:
Date From: 15 Jan 1912
Date To:
Verbatim Date:

RDE Sites Details
Country: United States of America
Province/State: New York

RDE Sites Details
County: Bronx Co.
City: Bronx
Precise Location: New York Botanical Garden. Cons. [Conservatory]
Township:
Range:
Section:
Latitude (Dec.):
Latitude (DMS):
Longitude (Dec.):
Longitude (DMS):
Elevation From: (m) (ft)
Elevation To: (m) (ft)

Determination RDE OCR Text Collection Features Habitat Specimen Processing Vernacular Assn

Edit Object 39 of 564 watson Admin 20085



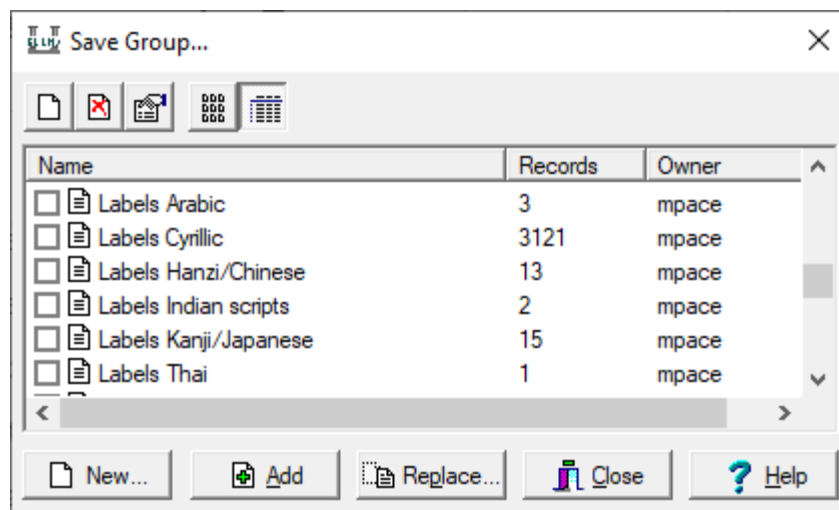
Ditto (F9)

Once all the information is entered, hit save and move on to the next record. If you have searched for all records from a particular collector, you could use F9 to Ditto the value from the previous record so you do not need to enter it again. F9 can be used in any field to Ditto the contents from the previously saved record.

Searching OCR and grouping records by similar values and using Ditto as much as possible is the best way to maximize the usefulness of RDE.

Symbols and Diacritics

With the exception of Lookup Lists, transcribe label data verbatim, in the original language, including symbols and diacritics. For labels with collection data in non-roman alphabets or character-based languages, please add the Catalogue record(s) to the appropriate Group. The topic of transliteration to Romanized text will be considered at a later date.



See here [how to enter symbols and diacritics](#).

See here a list of [common abbreviations](#) found on labels.

See here a digitized copy of [Botanical Latin by W. T. Stearn](#).

Specimens with 2 (or more) barcodes

Please keep in mind that OCR for sheets that have multiple labels will get included on the OCR tab for all specimens on the sheet. Please make sure to enter data on the RDE tab for that barcode only.

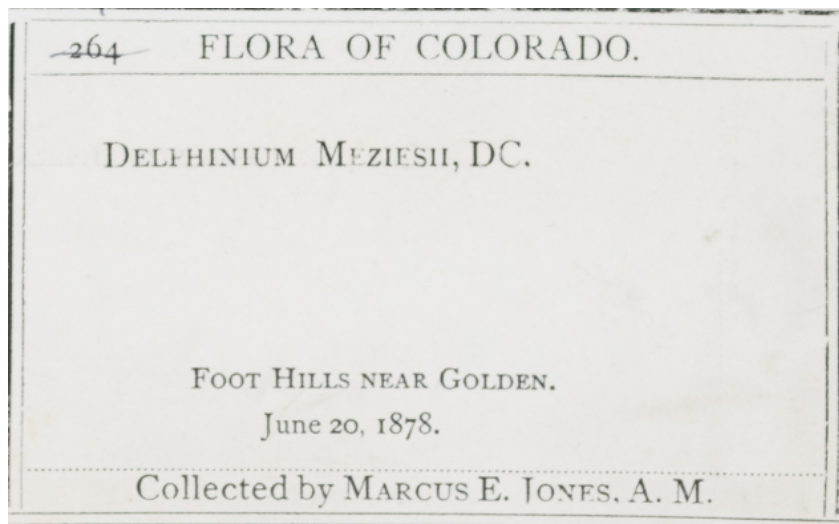
Buzzwords

It is recommended that you create a list of terms you find useful to group records. See Appendix 1 and 2 for a list of terms for identifying records from specific states for the MAM and NEVP projects. You can create similar lists for your own project. It is best to start a shared spreadsheet of buzzwords so all staff can use and add to the list. Please share the lists with your supervisors so we can create a master list for all projects.

Buzzwords should be any term or even common misspellings in the OCR text that helps find records related to your project.

Example:

To find records collected by Marcus E. Jones, it is easier to just search for “Marcus” and not use “Jones” in your search.



OCR output: “Marcus” is usually correct, while “Jones” is not even misspelled consistently

434 FLORA OF COLORADO.
Psoralea hypogaea, Nutt.
Colorado Springs; June 1879.
Altitude 5500 Feet.
Collected by Marcus E. Tontes, A. M.
02620820

FLORA OF COLORADO.
N ASTURTIU WOFFTICK ALE, R. Br.
Along the Piatte Rivo:
Denver, June 24, 1878.
Collected by Marcus E. Tones, A. M.
02377125

445 FLORA OF COLORADO.
Aconitum nasutum, Fisch. ss?:
Altitude, 8300 Feet.
Georgetown July 16, 1878.
Collected by Marcus E. Tontes, A. M.
Determined as:
ACONITUM 7
02727118

New York Botanical Garden
Hordeum pratense, Huds.
Altitude 6000 Feet.
Foot Hills near Golden, June so, 1878.
Collected by Marcus E. Jonks, A. M.
02121bS3

Ort. UL'101'11YLUJ-V
Phte. (Z*s*.+ie. Art'S* 2-/ Jj/V,
Drvro*. j
Between Colorado Springs, and Denver, June 6, 1878.
Collected by Marcus E. Tontes, A. M.
02785913

488 FLORA OF COLORADO
Myosurus minimum JL. 1
Colorado Springs. May 20, 1878.
Collected by Marcus E. Tontes, A. M.
02753287

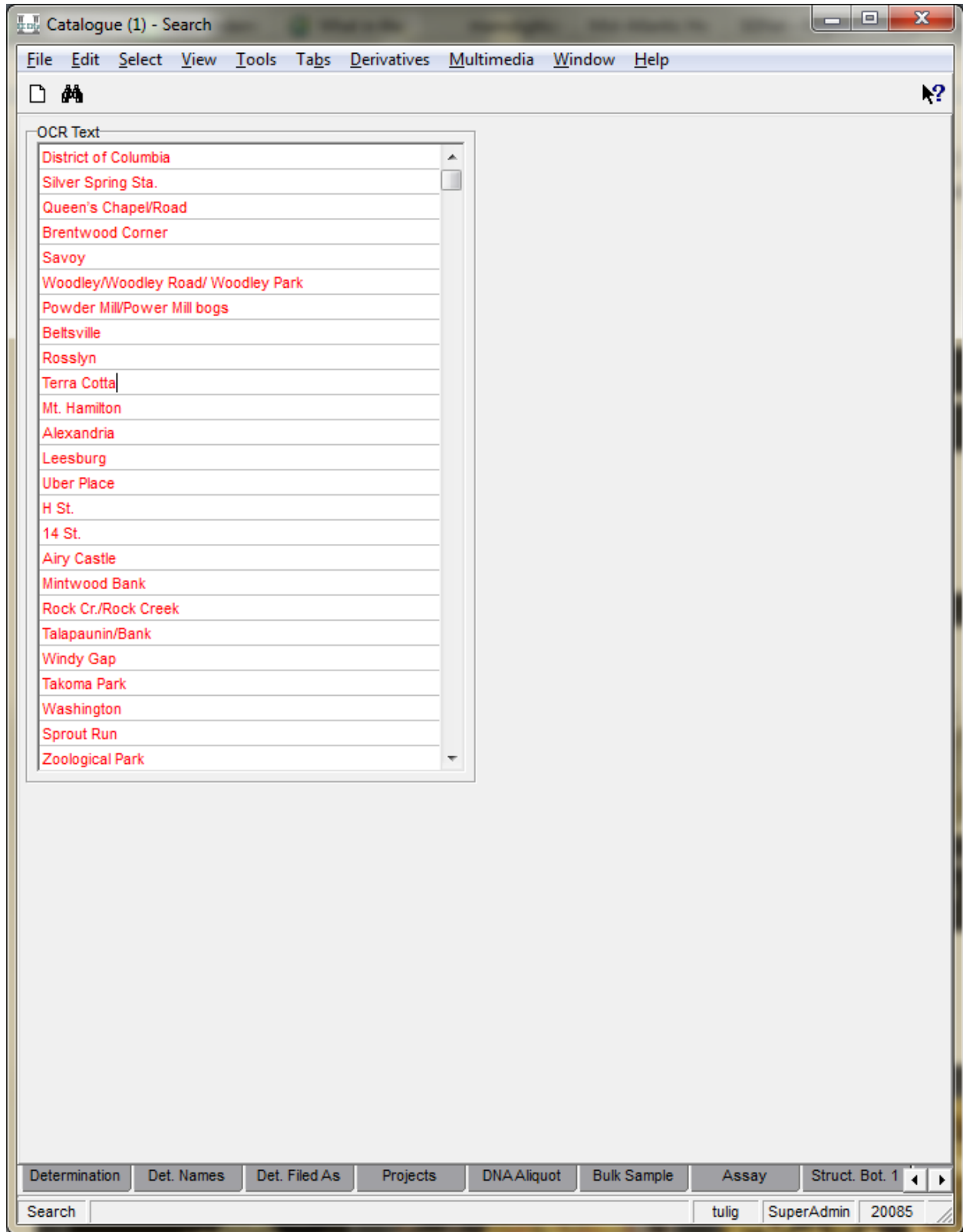
434 FLORA OF COLORADO.
Delphinium Meziesii, DC.
Foot Hills near Golden.
June 20, 1878.
Collected by Marcus E. Tontes, A. M.
02753287

Mass Search on Buzzwords

Since new OCR is added weekly, you will need to re-run your searches for the same buzzwords over and over again throughout your project. Rather than search for each word or phrase separately, it is easy to search for the whole list at once and see if any new records that fit your project's criteria have been added.

Copy and paste the list following these instructions:

[http://sweetgum.nybg.org/wiki/index.php/How to Cut and Paste a List into EMu](http://sweetgum.nybg.org/wiki/index.php/How_to_Cut_and_Paste_a_List_into_EMu)



If you get many results, keep starting again with individual terms to narrow your results. If you get few results, spend more time barcoding and imaging specimens to create a larger dataset.

Monthly Report

At the end of every day, please query for the number of RDE records you transcribed and enter the number of hours and the number of records transcribed on your monthly report under “Stage 2 – Complete Data Entry.”

If you still have RDE default query values loaded, change your query default values to “No defaults”

Query the Catalogue module via the Admin Tab by using the “Modified By” and “Modification Date” fields. Enter your full name and select the date by clicking on the calendar button.

(If your name does not appear in the Lookup List, you can still type in your first and last name to run the query. The Lookup List update is backlogged and not all names are appearing quickly. BUT the system still stores your name in each record.)

The screenshot shows a software window titled "Catalogue (1) - Search". The window has a menu bar with "File", "Edit", "Select", "View", "Tools", "Tabs", "Derivatives", "Multimedia", "Window", and "Help". Below the menu bar is a toolbar with icons for a folder, a person, and a help icon. The main area is divided into several sections:

- Internal Record Number:** A large empty text field.
- Summary Data:** A large empty text field.
- Legacy Data:** A large empty text field.
- Preferred GUID:** A section with a "Type:" label and a dropdown menu, and a "GUID:" label with a text field.
- GUIDs:** A section with "Preferred:" and "Type:" labels and dropdown menus, and a "GUID:" label with a text field.
- Insertion Details:** A section with "Inserted By:" (dropdown), "Insertion Date:" (calendar icon), and "Insertion Time:" (text field).
- Modification Details:** A section with "Modified By:" (dropdown showing "Melissa Tulig"), "Modification Date:" (calendar icon showing "15/03/2017"), and "Modification Time:" (text field).
- Import Details:** A section with "Import Identifier:" (text field) and "System Identifier:" (text field).

At the bottom of the window is a tabbed interface with tabs for "DwC 1.2 (2)", "Notes", "Multimedia", "Tasks", "Security", "Audit", and "Admin". The "Admin" tab is currently selected. Below the tabs is a search bar with the text "Search" and a status bar with the text "tulig Admin 20085".

If you did another task in EMu that day (like barcoding), then you will need to query for the number of records you modified that day, WITH an attachment in the RDE Collector field by entering \+ in that field. (To find just the number of records you barcoded that day, use the “Inserted By” and “Insertion Date” fields instead.)

Catalogue (1) - Search

File Edit Select View Tools Tabs Derivatives Multimedia Window Help

RDE Status
Processing Status:

RDE Collection Events Details
Collector: \+
Collection Number Prefix:
Collection Number:
Collection Number Suffix:
Collection Team:
Date From:
Date To:
Verbatim Date:

RDE Sites Details
Country:
Province/State:
County:
City:
Precise Location:
Township:
Range:
Section:
Latitude (Dec.):
Latitude (DMS):
Longitude (Dec.):
Longitude (DMS):
Elevation From (m):
Elevation From (ft):
Elevation To (m):
Elevation To (ft):

Projects DNA Aliquot Bulk Sample Assay Struct. Bot. 1 Struct. Bot. 2 Container Tissue

Search tulig Admin 20085

Data Entry Rates

After a few weeks of practice, staff have been able to average over 30-50 RDE records per hour. For records with limited data, rates can be as high as 70/hour.

Processing Status

Once a record is set to Processed, a Collection Events and Sites record will be automatically created with the values from the RDE tab and attached to the Catalogue record. Updating records to Processed is the responsibility of the information managers and can be left blank by all other users.

Once a record is marked Processed, do NOT update content on the RDE tab again. Make edits in the attached Collection Events or Sites modules.

RDE Cheatsheet

Summary

1. Log in to EMu as user RDE
2. Open Catalogue module
3. Load default values for your project (Edit->Default Values->Change->RDE default values->[your project here])
4. Use OCR text to find similar records to database (try to find groups of less than 10-20 records)
5. Use the RDE tab to enter collection data
6. Save record and move to the next record (F6)
7. Use ditto (F9) in any field exactly the same as the previous record
8. Repeat searches to find similar groups of records, use Previous search to add on to previous search or new search to start over
9. Change default values to No defaults, close catalogue (Edit->Default Values->Change->No Defaults)
10. Open catalogue module, record your data entry stats at the end of every day

Attaching a Collector is Mandatory. This prevents the record from showing up in future OCR searches.

Keyboard shortcuts

F5	Move to previous record
F6	Move to next record
F9	Ditto value from previous record
F12	Check value against look up list (enter a few characters, hit F12)
Ctrl+S	Save record
Tab	Move to next field
Shift+Tab	Move to previous field

What do I do if there is:

No collection number? Enter s.n. in the Collection Number Suffix field

No collection date? Enter s.d. in the Verbatim Date field

No collector? Attach Parties record: Collector unspecified

Illegible collector? Attach Parties record: "Illegible collector name" (if entirely illegible), or "see Collection Note" (if partially illegible, then add a note in "Collection Notes."

Other illegible text? Use the Collection Notes or Location Notes fields on the Collection tab; **do not make notes in RDE fields**. Place square brackets [] around your note, and quotation marks around any text that you transcribe verbatim within a note field.

Appendix 1

Example Buzzwords compiled by Ginger Apolo for the MAM project

❖ DC

- District of Columbia
- Silver Spring Sta.
- Queen's Chapel/Road
- Brentwood Corner
- Savoy
- Woodley/Woodley Road/ Woodley Park
- Powder Mill/Power Mill bogs
- Beltsville
- Rosslyn
- Terra Cotta
- Mt. Hamilton
- Alexandria
- Leesburg
- Uber Place
- H St.
- 14 St.
- Airy Castle
- Mintwood Bank
- Rock Cr./Rock Creek
- Talapaunin/Bank
- Windy Gap
- Takoma Park
- Washington
- Sprout Run
- Zoological Park
- Bennings
- Cailin Springs
- West Road
- Daniel's Road
- Brookland
- Chevy Chase
- Mt. Pleasant
- Zoological Park
- Deanwood
- Woodridge

❖ Delaware

- Bethesda
- Halltown Road (Road 50)
- Hourglass Road

- Town of Delmar
- Federalsburg Road
- Cannon Road
- Town of Cannon
- Elks Road (Road 46)
- Swain and Truitt Ditch
- Fenwick Island
- Atlantic Avenue
- Wilmington/near Wilmington
- Delaware River/RR
- Town of Milford
- Tub Mill Pond
- Town of Seaford
- Harrington Street
- Rehoboth/Rehoboth Beach
- Gravelly Branch
- “The Manual Range”
- Odessa
- Charleston
- Sparrow Bush/opposite Sparrow Bush
- Ellendale/Redden State Forest
- Rum Bridge Road (Road 484)
- Townsend
- Millsboro
- Bowers Beach
- Georgetown
- Gravel Hill
- Lewes Georgetown Highway (Road 18)
- Bethel
- Town of Concord
- Concord Pond Road (Road 516)
- Along/Phillips Landing Road (Road 496)
- Cow Bridge Ranch
- Steiner Road (Road 320)
- Midlands Wildlife Area
- Jones Store Road (Road 435A)
- Stockley Center
- Craigs Mill Pond

❖ Maryland

- Cabin John
- Federalsburg
- At edge of agricultural fields

- Idylwild/Wildlife Management Area
- Hidden Dream Training
- Solley Farm
- Patuxent River
- Dunkirk
- Ferry Landing
- Town of Finzel/Finzel
- Piney Run/Reservoir
- Crofton
- Lake Waterford Park
- Pasadena
- Nutter's Neck/Nanticoke River
- Kelly Ponds area of Patuxent River Park, Conway Road and Patuxent Road
- Odenton
- Chesapeake Bay/Beach
- Bethesda
- Town of Preston/Preston
- Kent Island
- Town of Chester/Chester
- Town of Fruitland/Fruitland
- Town of Arnold/Arnold
- Ritchie Highway
- Church Lane
- Queen Anne Highway
- Bush River
- Hayman Drive
- Federalsburg Highway
- Saint Michaels Road
- Caroline Industrial Park
- Caroline Drive
- Frostburg
- Snow Hill
- Havre de Grace
- Latimer's Point
- VFW building

❖ New Jersey

- Metedeconk/River
- Delaware River
- Milford
- Tom's River
- Point Pleasant

- Carlstadt
- Washington Township
- Town of Chatsworth
- Tabernacle Chatsworth Road
- Mount Ephraim
- Newton Creek
- Browntown
- Watchung
- Closter
- Berkeley/Heights
- Great Swamp
- Woodbury
- High Line
- Budd's Lake
- Second Mountain/2nd Mountain/2nd Mt.
- Eagle Rock
- Seeley's Notch
- Scotch Plains
- Cape May/Cape May Court House
- Forked River
- Culver's Lake
- Metuchen
- South River
- Great Swamp
- South Amboy
- South Brunswick
- Deans Rhode Hall Road (CR 610)
- Flemming Pike
- Hammonton
- North Plainfield
- Palatine
- Freehold
- Como
- Somer's Point
- May's Landing

❖ Philadelphia

- Lower Wissahickon Valley
- Darby Creek
- The Willow Grove
- Lost Creek Gap
- McAllisterville

- Along/Hammer Creek
- Brickerville
- North White Deer Ridge
- South/Williamsport
- Bald Mountain
- Scranton
- Windfall Road
- Fleetville
- Little Pine Creek
- English Center
- Little Tacony Creek
- Melrose
- Altoona
- Clearfield
- Clearfield Creek
- Dodge Hill
- Eaglesmere
- Rock Hill
- Hopewell Furnace/Village National Historic Site
- Along/Youghiogeny River
- Ohio Pyle
- Barkleyville
- Wesley
- Highway/Hwy. 8
- Trythall
- Greenhill
- Allen's Lake
- Barclay Mountain
- Franklindale
- Meadville
- Woodcock Township
- Fisher's Ferry
- The Welsh Mountains
- Churchtown
- Beartown
- Conemaugh-Jenner Plateau
- The Allegheny Mountains

Appendix 2

Example Buzzwords compiled by Natalie Feldsine for the NEVP project

- NY Queries
 - localities

- \”long island\”
- \”staten island\”
- \”new York botanical garden\” \”local herbarium\”
 - This is a particularly good query as it gives you many of the NY or NJ specimens that have handwritten labels
 - \”local herbarium\” without the New York qualifier also brings up many specimens
- \”new York botanical garden\” \”cultivated plants\”
- \”vicinity of new York\”
- State abbreviations ie. \”conn.\” \”mass.\” \”V.T.\” etc. (isn’t always effective for two letter abbreviations)
- Bronx
- \”flora of new York\”
- \”plants of new York\”
- \”marthas vineyard\” \”martha’s vineyard\”
- \”Nantucket\”
- Catskills, catskill
- \”mt. Washington\” \”mount Washington\” (place in NH and in Massachusetts)
- Adirondack, Adirondacks
- \”northern new York\”
- \”cold spring harbor\”
- “white mountains” (may need to specify not New Mexico, Arizona, or California to narrow it down to NH specimens)
- \”plants of\” Hampshire \!”west Virginia\”
 - Brings up some New Hampshire specimens when OCR doesn’t pick up the “New”
- Maine counties/rivers
 - Penobscot
 - Aroostook
 - Piscataquis
 - Sagadahoc
 - Androscoggin
- NH localities
 - Ossipee
- Berkshire
- \”lake champlain\”
- \”central new york\”
- Cayuga
- \”herbarium n.y. state museum\”
- Yonkers
- \”new London\” in CT
- \”penn yan\” brings up H.P. Sartwell’s local specimens
- \”USA : NY\”
- \”, N.Y.\” \”co., N.Y.\” or \”county, new York\” will bring up specimens where the county name may be handwritten or not picked up by OCR.

- \", N.Y.\" brings up a lot of other specimens that can be filtered through by reading the OCR text. This query template can also be used for other state abbreviations
- \\"tri-state\"
- o Collectors
 - \\"w.c. ferguson\"
 - \\"Herbarium of William Cashman Ferguson\"
 - \\"Eugene Pintard Bicknell\"
 - \\"Victoria Bustamante\" (several NY and Canada specimens)
 - \\"j. w. congdon\" collects in MA, RI, and CA
 - \\"r. keir\" – Maine and Canada specimens
 - \\"willard w. eggleston\" or just \\"eggleston\"
 - He collected everywhere but querying his name helps to locate NEVP specimens that do not show up in other OCR queries. Often OCR only picks up a partial word or the words on the label are distorted in the OCR text. In order to filter, I suggest clicking on the OCR tab to read the OCR text and look for NEVP specimens.
 - \\"Harold n. moldenke\"
 - \\"m. r. north\" collects in CT
 - \\"frank c. mackeever\"
 - \\"ferdinand Blanchard\" collects often in VT
 - \\"N.L. Britton\" brings up MANY specimens that have handwritten labels (although you will have to filter through some specimens)