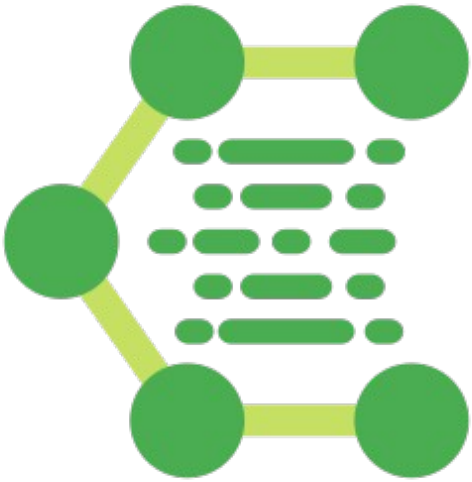
An Introduction to collNotes and collBook



Caleb Powell
CalebadamPowell@gmail.com

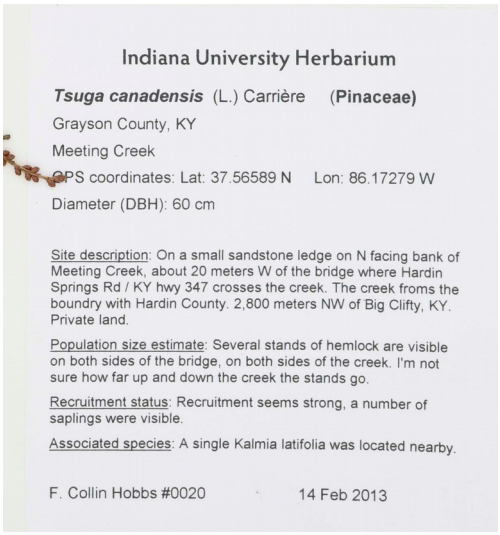
Jacob Motley
programmingisfunjacmot@gmail.com

Updated: 2019-08-24

Table of Contents

Introduction - “Born Digital” Natural History

Digitized natural history collections provide millions of biodiversity records to online public portals such as GBIF[1], iDigBio[2], and SERNEC[3]. These datasets are important resources for researchers and decision makers, but keeping them up-to-date is an ongoing challenge. Much of the data associated with natural history records is initially stored as text on a specimen label (Fig. 1) which need to be transcribed for use in online portals.


Indiana University Herbarium

Tsuga canadensis (L.) Carrière (**Pinaceae**)

Grayson County, KY

Meeting Creek

GPS coordinates: Lat: 37.56589 N Lon: 86.17279 W

Diameter (DBH): 60 cm

Site description: On a small sandstone ledge on N facing bank of Meeting Creek, about 20 meters W of the bridge where Hardin Springs Rd / KY hwy 347 crosses the creek. The creek froms the boundry with Hardin County. 2,800 meters NW of Big Clifty, KY. Private land.

Population size estimate: Several stands of hemlock are visible on both sides of the bridge, on both sides of the creek. I'm not sure how far up and down the creek the stands go.

Recruitment status: Recruitment seems strong, a number of saplings were visible.

Associated species: A single Kalmia latifolia was located nearby.

F. Collin Hobbs #0020 14 Feb 2013

Fig. 1: A data rich specimen label from the Indiana University Herbarium (catalog number: IND-0006820).

collNotes and collBook have been developed to accomodate collecting “born digital” *Fungi* and *Plantae* records, which are initially gathered in digital formats suitable for online portals. To achieve this, collNotes, a mobile application, was developed to supplement the traditional field journal. A companion desktop application, collBook, enables users to refine field notes into print ready lables and portal ready comma separated value (CSV) files. Both programs have been released as open source projects[4, 5], in hopes that continued feedback and contributions from the community will improve the tools. For a more detailed overview of collNotes & collBook, see the associated publication[6].

1 Global Biodiversity Information Facility: https://www.gbif.org
2 Integrated Digitized Biocollections: https://www.idigbio.org/portal
3 SouthEast Regional Network of Expertise and Collections: http://sernecportal.org/portal
4 collBook’s primary repository: https://github.com/CapPow/collBook
5 collNote’s primary repository: https://github.com/j-h-m/collNotes
6 https://doi.org/10.1002/aps3.11284

Getting Started

collNotes (mobile) Installation

collNotes is a free mobile application on Android and iOS. The applications can be located in the appropriate app stores [7,8] or by following the included QR code (Fig. 2). Since GPS functionality is not dependent on cellular service, collNotes does not require a data enabled device. In addition to personal cellular devices, collNotes can also be used on a tablet.

Fig. 2: QR code linking to mobile application.

collNotes (mobile) Initial Setup

Preferences for “default primary collector name” and “starting field number” are available in the Settings menu (Fig. 3) which is accessible from the main menu. Before using collNotes in the field it is highly reccomended to test and become familiar with the program. Tests should include generating and exporting site and specimen level records. It is also reccomended to review the “Making Site Level Observations,” and “Making Specimen Level Observations” sections of this document.

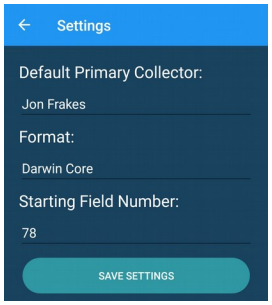


Fig. 3: collNotes settings

collBook (desktop) Installation

collBook is a free desktop application available for Linux, OS X, and Windows[9]. The program is written, and can be run in Python. Additionaly, executable binaries are available for OS X and Windows. When run as an executable, the program does not need to be “installed” in a formal sense, it simply runs from the executable program.

7 Google Play store link: https://play.google.com/store/apps/details?id=com.thc.collnotes
8 iOS App store link: https://itunes.apple.com/US/app/id1455049130?mt=8
9 CollBook installation instructions: https://github.com/CapPow/collBook#installation

collBook (desktop) Initial Setup

collBook is a feature rich program with many customization options. The “Preferences” menu is accessed using the gear icon on the tool bar (Fig. 4). Multiple categories are selectable along the left frame of the Preferences window. Customizing settings in the “Taxonomy,” and “Labels” categories is reccomended. For example, in the “Taxonomy” category there are settings for name and authority change policies. These policies dictate how taxonomic alignments should be handled.

collBook (desktop) Overview

Before using collBook, it is highly reccomended to test and become familiar with the program. Tests should include importing, refining, and exporting records. Each of these functions are available on the toolbar. Hovering the mouse over toolbar icons will display a brief description of the function. The collBook user interface contains four panes: a label preview, form view, site navigator, and table view (Fig. 4).

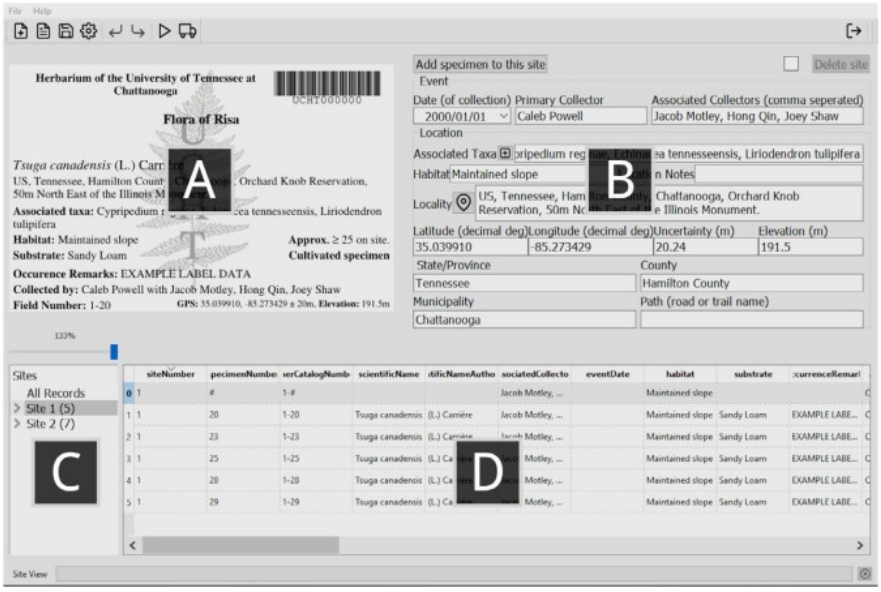


Fig. 4: The collBook user interface, showing (A) the label preview, which presents an image of the label to be produced for a selected record; (B) the form view, where data pertinent to the selected class (e.g., “Site 1”) may be edited; (C) the site navigator, which is used to select record(s) for editing or refining; and (D) the table view, which provides an overview of the selected record(s) and allows unrestricted edits to the data.

Using collNotes & collBook

Data Organization

Records are stored in collNotes & collBook in a hierachal data structure, which can be thought of as a series of parent-child relationships where data from parent records are inherited (copied) to child records. This organization reduces the need to enter redundant information. In collNotes, this hierachal structure is reflected in nested menus where, for example, site records may only be created after creating or selecting the trip they are associated with. In collBook, this hierachal structure is reflected in the way records (and groups of records) are selected in the site navigator (Fig. 4C). Edits made in either collNotes, or collBook will be inherited by the appropriate records. For example, in collBook when a site is selected edits made using the form view (Fig. 4B) are applied to every specimen collected from that site.

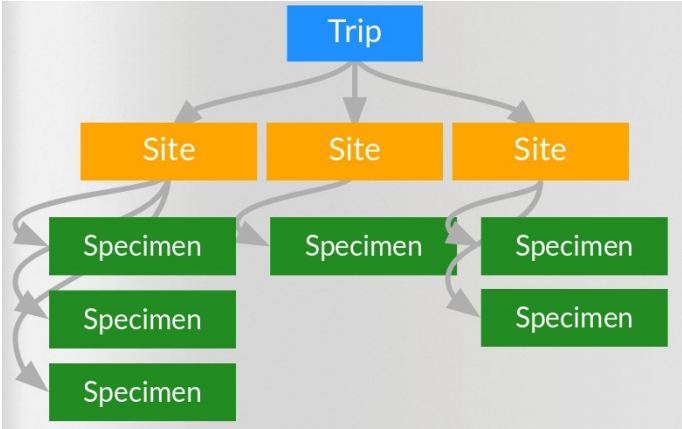


Fig. 5: A visualization of the hierachal data structure used in collNotes. In collBook, trip level data is called "All Records."

Making Site Level Observations

When making site level observations, the geographic size of a single site is left up to the researcher's descretion. Since site data such as locality, habitat, and GPS coordinates, are inherited by each specimen collected from a single site, it is generally reccomended to keep site ranges between 5 – 30 meters. New sites are always appropriate when there are major habitat changes. For example, specimens gathered from a cliff face and a creek bed 5 meters apart should be from seperate sites.

Locality data recorded for site level observations will be refined in collBook using the Google's Reverse GeoLocation service. The result is a textual preamble based on the GPS coordinates which is added to the locality string recorded with collNotes in the field. For example, the locality string: "50m northeast of the Illinois Monument." would become: "US, Tennessee, Hamilton County, Chattanooga, Orchard Knob Reservation, near East 4th Street, 50m North East of the Illinois Monument."

Making Specimen Level Observations

When specimen records are generated, they will include a field number structured as a "site number - specimen number". This number should be used to maintain the connection between the physical specimen and the digital record. In collNotes the field number is displayed on the top of the screen when generating a specimen level record. For example, in Fig. 6, the field number is "46-79" signifying site 46, specimen 79. In collBook the field number is displayed on both the preview label as well as in the site selection pane (Fig. 4).

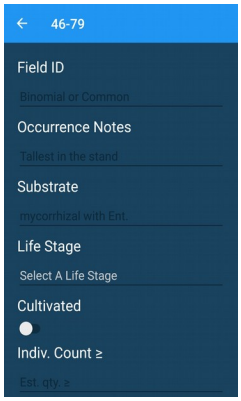


Fig. 6: collNote's Specimen level record generation menu.

Specimen specific data may be gathered in the field using collNotes or generated while making formal identifications using collBook. Recording specimen specific observations in the field using collNotes can produce more detail rich records (e.g., containing abundance estimates) but requires more time in the field. In some cases a researcher may prefer to record only site level observations and store multiple taxa from the same site together, so called "hay baleing." In this second scenerio, specimen specific records are generated in collBook while formal identifications are being made. When this method is used, the researcher should ensure the site number is labeled on the container to maintain the connection between the physical specimens and the digital record.

Exporting Records from collNotes

Records in collNotes are stored locally on the mobile device until they are exported. Field conditions can be unpredictable and if a device is lost or destroyed un-exported records will be lost. Exporting records from collNotes does not delete data from the device, it sends a copy of the data to the chosen destination. So, it is reccomended to export records as often as is convienent and reasonable.

When exporting records from collNotes, a UTF-8 encoded .CSV file is generated and the mobile device's share settings are invoked to decide how to transfer that file. Many users find it convienent to simply E-mail it to themselves, which does require wifi or cellular data service. The .CSV file generated by exporting records will be organized (with few exceptions) under The Darwin Core terms[10]. Exported record files will have rows for both specimen, and site specific data.

Using collBook without using collNotes

collBook has been written to be interoperable with collection protocols which do not include collNotes. collBook can load data gathered in Maya-Lastra's field data collection application "ColectoR,"[11] or from the iNaturalist based workflow proposed by Heberling and Isaac[12]. Since these sources do not incorperate a hierachal data structure, collBook assigns a unique site for every specimen loaded this way. Records may also be manually transcribed into collBook.

10 The Darwin Core quick reference guide: http://rs.tdwg.org/dwc/terms
11 Maya-Lastra, C. A. (2016), https://doi.org/10.3732/apps.1600035
12 Heberling, J. M., and B. L. Isaac. (2018), https://doi.org/10.1002/aps3.1193

Refining Records in collBook

The “play” icon on the collBook toolbar is used to refine records. This feature will **only refine** those records which are **currently selected.** The “All Records” option in the site navigator (Fig. 5C) will select all records. Refinements are performed in this order:

- Taxonomic verification, to verify name status and fill in blank or incorrect authorities. The source and behaviour is customizable in preferences.
- Reverse geolocation, to populate location data from GPS coordinates.
- Associated taxa assignment, to assemble a list of taxa identified from the same site (optional, based on preferences).
- Catalog number assignment, and barcode generation (optional, based on preferences).

Saving a Working File in collBook

While working with records in collBook which are not yet ready for exporting, it is possible to save a working file which can be opened later. The “diskette” icon on the collBook toolbar is used to save a working file. This working file contains collNotes / collBook specific data and is not the same as the portal ready .CSV file produced when exporting records from collBook.

Exporting Records from collBook

After refining records in collBook, it is reccomended to review each record’s label preview (Fig. 5A) for accuracy. The “truck” icon on the collBook toolbar is used to export records for database upload and label printing. This feature will **only export** those records which are **currently selected.** The “All Records” option in the site navigator (Fig. 5C) will select all records. Exporting will produce two files, a UTF- 8 encoded .CSV file organized (with few exceptions) under The Darwin Core terms[13], and a .PDF file with ready-to-print labels.

13 The Darwin Core quick reference guide: http://rs.tdwg.org/dwc/terms

Frequently Asked Questions (FAQ)

Q: I would like to see " *x,y,z* " feature included in collNotes / collBook, will you add it?
A: Great! We're activly seeking community feedback, please post your feature request to collBook's issue page (https://github.com/CapPow/collBook/issues).

Q: I have discovered a bug in collNotes / collBook, how do I report it?
A: It would be very helpful if you would post a bug report on collBook's issue page (https://github.com/CapPow/collBook/issues). In addition, if collNotes crashes and your mobile device asks if you'd like to send a crash report, please do so! We see those reports and they are helpful.

Q: How can I contribute to the project (collNotes, collBook, or both)?
A: Amazing! We are activly seeking additional contributors. Pull requests are welcomed for both collNotes and collBook. If you have any problems, or want to discuss collaboration please contact us.

Q: How can I delete data from collNotes?
A: You cannot, yet. This is actually a known issue which is being addressed in an upcoming update to collNotes. You can check the progress here.

Q: What does coll in collNotes / collBook stand for?
A: It refers to the word "collection."

Q: How should I cite collNotes / collBook?
A: Please cite the associated publication: https://doi.org/10.1002/aps3.11284

Thanks for reading this far!