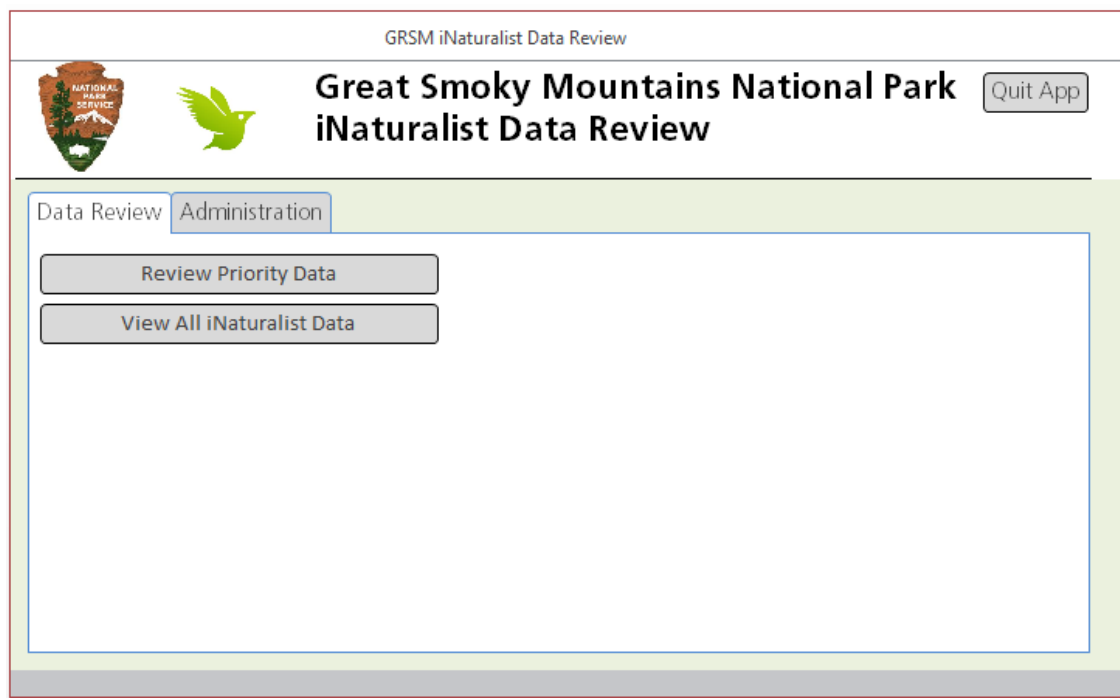


iNaturalist Data: Soft Integration with GRSM Data Processes and Repositories

Contents

Background.....	2
Soft Integration.....	2
Downloading iNaturalist Data	2
Importing iNaturalist Data into the Data Processing App.....	3
Targeted Species.....	4
Reviewing Observations for Quality.....	5
Processing Validated Data.....	5
Obscured Species Observations.....	6



Background

iNaturalist is an 'online social network of people sharing biodiversity information to help each other learn about nature' (iNaturalist website). While iNaturalist contains a wealth of data, the crowd-sourced nature of the data has made it challenging to integrate into GRSM data repositories on any wholesale level. While many of the observations are likely valid, it is challenging to know identify these high quality records.

Soft Integration

Instead of ingesting all, or some, of the iNaturalist observation into our systems, such as NPSpecies, we propose a 'soft integration' methodology. Soft integration means that we will use these data as an early warning system for encroaching and/or invasive species, an alert to potentially new location of species of special concern such as green-headed coneflower (which is undergoing review as collection by Eastern Band of Cherokee Indians is planned for), a source of data for high-importance species such as species for which we only need a few more records to be able to utilize in the Atlas of the Smokies applications which require at least 30 spatially distinct observations.

Downloading iNaturalist Data

- iNaturalist manages 'places' (including [GRSM](#)) which are areas with a specific boundary—all observation within that boundary are associated, regardless of being within a project or not
- Access 'places' from the 'More' drop down on the home page of iNaturalist



Figure 2. You can access observations associated with a place by clicking 'Observations' tab on the place-specific, tabbed menu

The place-specific data can be downloaded using the export observations page:

<https://www.inaturalist.org/observations/export>

- Query parameters:
 - With photos
 - Place: Great Smoky Mountains National Park, NC, US
 - Date Range: populate start date with start date= to date of last export
 - Include all taxonomic levels

Those query parameters generate the following search text in iNaturalist. You can edit the search text and paste it into the search bar.

search_on=place&has%5B%5D=photos&quality_grade=any&identifications=any&place_id=72645
&d1=2018-10-02&d2=2019-04-11

Export Observations

1 Create a Query

Create an observation query just like you would elsewhere on the site. You can also cut and paste an observations URL from another part of the site. You must specify a **taxon**, **place**, **user**, **project**, or **search query**.

Figure 3. The query interface in iNaturalist allows you to paste search text into the search bar.

Importing iNaturalist Data into the Data Processing App

Considerations:

- Observed_on- must be imported as a short text field instead of a date field because of the way the date is formatted
- Description- must be imported as long text instead of short text to avoid field truncation

The following fields were added to the export to facilitate data review, validation, and export to Evidence Template for inclusion in NPSpecies.

- Flag (Short Text)- value from tlu_Species that denotes why the species is a priority for review
- DataStewardRecordReview (Yes/No)- Record reviewed by data steward
- ValidRecord (Yes/No)- Used to mark whether or not the record is valid according to the data steward, only record with ValidRecord=Yes will be imported into NPSpecies Evidence
- DataStewardReviewer (Short Text)- Name of data steward who reviewed record
- UpdatedDate Date/Time)- Date record was updated/reviewed
- ExportedToEvidenceTemplate (Yes/No)- Record was exported to EvidenceTemplate
- DateExportedToEvidenceTemplate (Date/Time)- Date record was exported to Evidence Template

The following queries will be run on newly imported iNaturalist data.

- qup_ExportProcessing- this query replaces null values in the iNaturalist field called taxon_species_name with a value of "spp." to enable matching to our NPSpecies park species list.
- qup_UpdateFlagValues- this query adds the flag reason from the tlu_Species table to the iNaturalistExport table

- UPDATE tlu_FlagReasons INNER JOIN (INaturalistExport INNER JOIN tlu_Species ON INaturalistExport.scientific_name = tlu_Species.ScientificName) ON tlu_FlagReasons.Flag = tlu_Species.FlagReason SET INaturalistExport.Flag = [tlu_FlagReasons]![Flag];

Targeted Species

Connecting observations to meaningful species attributes allows us to utilize the records for various reason; delving into records of particular interest. For example, invasive plant species are one of the list of species that have been accounted for in this application. iNaturalist observations of invasive plant species can serve as an early warning system to resource managers.

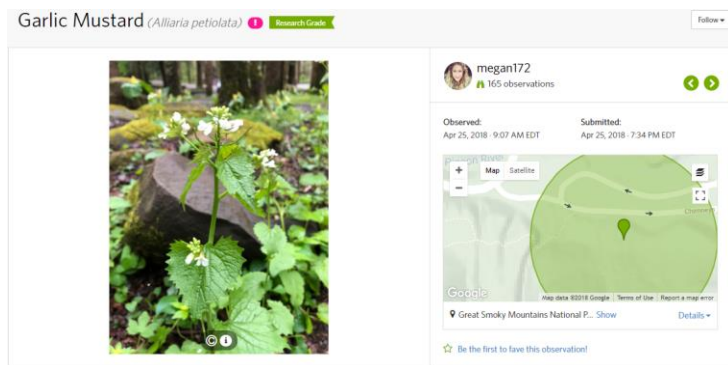


Figure 4. An iNaturalist observation record for garlic mustard, a category 1 exotic plant. Using the mapping interface can allow the data validator to see where the observation was on the landscape, perhaps cross referencing it with treatment sites or other relevant information.

Discover Life in America (DLIA), the park's non-profit partner on the All Taxa Biodiversity Inventory, has an iNaturalist project called [Species SnapIt & MapIt](#) lists of species that are relatively easy to identify and which need a few more observations in order to be included in the SpeciesMapper application. These lists have been turned into an iNaturalist guide which provides the list of priority species by taxa group as well as images of each species to aid in identification. The guide can also be used as a citizen science communication tool, informing Discover Life in America volunteers as well as the public at large about species observations that we are specifically targeting.

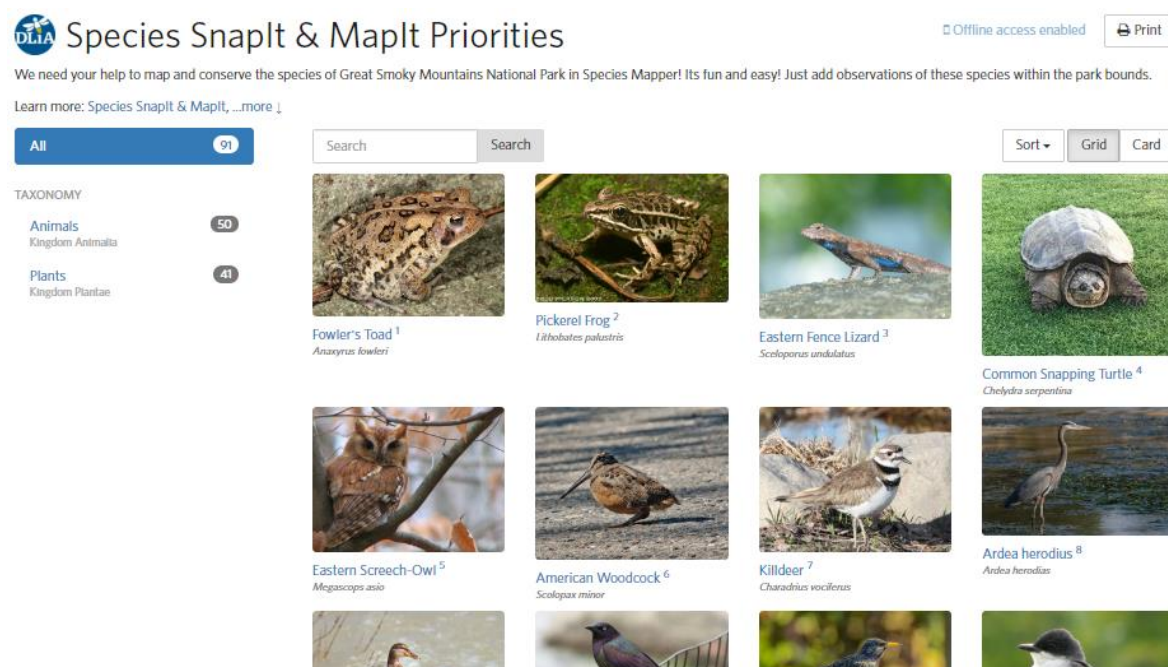


Figure 5. The Species SnapIt & MapIt guide lists our current priority species list.

Observations of these priority species are an important enough target to warrant manual review. On some interval, DLIA and GRSM will work together to update these lists: removing species which have enough observations to model on, and adding species as we see fit. Keeping these lists up to date is of the utmost importance in terms of volunteer engagement as well as validity of the soft integration workflow which is designed to be nimble and flexible depending on our needs and goals.

Similar to the processing of exotic plant observations (Figure #), priority species observations will be reviewed manually for quality. Different priority groups may be assigned to different data stewards depending on expertise.

Reviewing Observations for Quality

Records that are manually vetted require several considerations:

- Image review—does the image show the species clearly enough for validation of the identification?
- Location—the map viewer in the iNaturalist in user interface will serve as a basic way to visualize the coordinates to check for validity—would this species occur in this area of the park?
- Duplicate Records—if the combination of species, location and date indicate a duplication, choose one of the records to mark as valid. This kind of duplication can occur when two iNaturalist users observe the same individual species at the same time and location.
- Date—time of year can often be a deciding factor on whether or not the record is valid. If seasonality of the species is illogical, it can indicate that the species is misidentified.

iNaturalist Observations of Priority Species								
* Double-click on the field label to change sort order. Click the button next to each record to open the observation in iNaturalist.								
Close								
Filters								
Flag:	Scientific Name:	Common Name:	Filter Is On					
Exo-cat 1								
Flag*	Sci Name*	Common Name*	Obs Date*		Reviewed*	Reviewer*	Updated Date*	Valid Record*
Exo-cat 1	Microstegium vimineum	Japanese stiltgrass	9/21/2018		☐			☐
Exo-cat 1	Lonicera japonica	Japanese honeysuckle	6/7/2018		☐			☐
Exo-cat 1	Alliaria petiolata	garlic mustard	4/25/2018		☐			☐
Exo-cat 1	Lonicera japonica	Japanese honeysuckle	3/28/2018		☐			☐
Exo-cat 1	Lonicera maackii	Amur honeysuckle	4/22/2011		☐			☐
Exo-cat 1	Lonicera japonica	Japanese honeysuckle	10/7/2017		☐			☐
Exo-cat 1	Lonicera maackii	Amur honeysuckle	7/27/2017		☐			☐
Exo-cat 1	Hedera helix	common ivy	9/19/2016		☐			☐
Exo-cat 1	Microstegium vimineum	Japanese stiltgrass	5/13/2016		☐			☐

Figure 6. The data gateway allows the user to filter on flag, scientific name, and common name. The end user can open the iNaturalist record by clicking the globe button. Upon validation, the user should click reviewed, enter their name, the date, and then check valid if the result should be included in an NPSpecies import. Only records that are flagged as valid will be imported into NPSpecies Evidence.

Processing Validated Data

Observations that have been manually vetted will be processed for inclusion in NPSpecies by reformatting data into the Evidence Template (a structured Microsoft Access database that allows for bulk import into NPSpecies).

- qapp_FormatForEvidenceTemplate- this query reformats the vetted iNaturalist data into the Evidence Template which can then be imported into NPSpecies Evidence.

```
INSERT INTO tblObservations ( ParkCode, ObservationRecordID, TSN, TaxonCode, Sensitivity,
DocLatinName, [Date], DatePrecision, Location, latitude, longitude, LocationError, Comments,
DataSource, RecordStatus, InParkBoundary, EnteredBy, Project, OnlineLink,
ObservationMethodDetails )
SELECT "GRSM" AS ParkCode, INaturalistExport.ID AS ObservationRecordID, NPSpeciesList.TSN,
NPSpeciesList.[Taxon Code], 3 AS Sensitivity, INaturalistExport.scientific_name AS DocLatinName,
```

```
INaturalistExport.observed_on AS [Date], 1 AS DatePrecision, INaturalistExport.place_guess AS
Location, INaturalistExport.latitude, INaturalistExport.longitude,
INaturalistExport.positional_accuracy AS LocationError, "Reviewed by GRSM Staff Member: " &
[DataStewardReviewer] & "; Reviewed On: " & [UpdatedDate] & "; iNaturalist User_ID: " & [user_id]
& "; iNaturalist user_login: " & [user_login] AS Comments, "Record was obtained from iNaturalist
(naturalist.org). Record was reviewed by GRSM Staff and formatted for NPSpecies Evidence." AS
DataSource, 1 AS RecordStatus, 1 AS InParkBoundary, "Kendra Straub" AS EnteredBy,
"Project:iNaturalist Data Processing. IRMA Data Store Project Reference Code: 2259992" AS
Project, INaturalistExport.url AS OnlineLink, "Image used to verify species observation: " &
[image_url] AS ObservationMethodDetails
FROM INaturalistExport LEFT JOIN NPSpeciesList ON INaturalistExport.scientific_name =
NPSpeciesList.[Scientific Name]
WHERE (((INaturalistExport.ValidRecord)=True));
```

- qup_ExpoortMetadata- this query updates the iNaturalistExport table with export metadata
 - UPDATE INaturalistExport SET INaturalistExport.ExportedToEvidenceTemplate = True,
INaturalistExport.DateExportedToEvidenceTemplate = Now()
WHERE (((INaturalistExport.ValidRecord)=True));

Obscured Species Observations

From iNaturalist: 'To protect our users' privacy, the iNaturalist platform will not grant any organization or user access to private coordinates unless the observer has agreed to share them with you. The best way to do this is to create a Traditional project and reach out to the observers, asking them to join your project and add the observation(s) in question to the project. More information on Traditional projects can be found on our [Managing Projects](#) page.'

References

List of iNaturalist Observation fields: https://www.inaturalist.org/observation_fields/