

Darwin Core Survey Starter Guide (U.S. Federal Agency Version)

Why Use Darwin Core?

Darwin Core (DwC) is a community standard for sharing biodiversity data. Using DwC ensures that survey data are Findable, Accessible, Interoperable, and Reusable (FAIR) and compatible with global aggregators like GBIF. For U.S. federal agencies, DwC aligns with Project Open Data principles and supports open science.

1. Minimum Darwin Core Fields

At a minimum, every biodiversity survey should include these core fields:

| Field (Darwin Core) | Plain Language Description | Example |
|---------------------|--|------------------|
| scientificName | The scientific name of the organism | Quercus alba |
| eventDate | The date of the observation or collection | 2023-08-14 |
| decimalLatitude | Latitude in decimal degrees (WGS84) | 38.8951 |
| decimalLongitude | Longitude in decimal degrees (WGS84) | -77.0364 |
| recordedBy | Name(s) of person(s) who recorded the data | Jane Smith |
| basisOfRecord | Type of record (e.g., HumanObservation, PreservedSpecimen) | HumanObservation |
| occurrenceID | A unique ID for each record | USFS-OBS-0001 |

2. Example Data Table (CSV)

```
occurrenceID,scientificName,eventDate,decimalLatitude,decimalLongitude,recordedBy,basisOfRecord,individualCount,samplingProtocol
USFS-OBS-0001,Quercus alba,2023-08-14,38.8951,-77.0364,Jane Smith,HumanObservation,3,Visual survey
USFS-OBS-0002,Cardinalis cardinalis,2023-08-14,38.8951,-77.0364,Jane Smith,HumanObservation,1,Point count
```

3. Metadata (README.md)

```
# [Project Name] Biodiversity Survey

## Description
This dataset contains observations collected during [survey/project name]. Data follow the Darwin Core

## Geographic Coverage

- Location: [Region, State, Coordinates]
- Coordinate system: WGS84

## Temporal Coverage
- Start date: YYYY-MM-DD
```

- End date: YYYY-MM-DD

Methods

- Sampling protocol: [Describe methods used]
- Effort: [Number of plots, transects, hours, etc.]

Contact

- Name: [Data steward / PI]
- Organization: [Agency name]
- Email: [contact@example.gov]

License

This dataset is released under CC0 1.0 (recommended for federal datasets).

4. Suggested GitHub Repository Structure

```
/data
  survey_data.csv
/docs
  README.md
LICENSE
```

5. Quality Checks

- Dates are in ISO format: YYYY-MM-DD. - Coordinates are decimal degrees (lat: -90 to 90, lon: -180 to 180). - No personally identifiable information (PII) included. - Each record has a unique occurrenceID. Optional: validate with GBIF Data Validator (<https://www.gbif.org/tools/data-validator>).

6. Publishing Your Dataset

1. Create a free GitHub account. 2. Create a new repository (e.g., forest-bird-survey-2023). 3. Upload survey_data.csv, README.md, and LICENSE. 4. Share the repository link with colleagues and collaborators.

7. References & Resources

- Darwin Core Quick Reference Guide: <https://dwc.tdwg.org/terms/> - GBIF Darwin Core Validator: <https://www.gbif.org/tools/data-validator> - Project Open Data Metadata: <https://project-open-data.cio.gov/metadata/> - Creative Commons Licenses: <https://creativecommons.org/share-your-work/>