

Species Data Deliverable Specifications

Background

For the NPS Inventory and Monitoring Division Inventory and Assessment Program (I&A), data products should be in open, non-proprietary formats and include sufficient metadata and supporting documentation. Example templates are comprehensive and can be extended as appropriate for specific projects. A data acceptance workflow is in place in which submitted data are screened for protected information and for adherence to deliverable requirements.

General Guidelines

- Datasets are submitted in delimited format containing plot- and/or point-level evidence of occurrence or lack of occurrence (non-detects) of focal taxa; habitat conditions including the relative abundance of broad land-cover classes of habitat; and variables used in determining effective survey effort and detection probability of the focal taxa. Project-specific elements may be added as needed; include element descriptions. If you do not include Ecological Metadata Language ([EML](#)) format metadata, this will be created on your behalf.
- Spatial data layers are submitted in an open format (e.g., GeoPackage/SQLite) documenting plot polygons and transects with Ecological Metadata Language-format or FDGC format metadata.

Templates and Tools

- [Spreadsheet template](#) (Microsoft Excel or comma-delimited) with Event, Occurrence, and Taxon elements.
- [ezEML](#): An intuitive, online form-based EML editor created and maintained by the Environmental Data Initiative (EDI). ezEML is not suitable for data packages that contain Controlled Unclassified Information (sensitive) because it requires that you upload your data to an external server. This [video](#) gives an overview of the ezEML features.
- [EML metadata template](#): includes spatial and taxonomy components; at minimum, include the bounding coordinates for the park/project area, sample plot centroids, and focal taxons (taxonomic domain used).

- EML R package: <https://cran.r-project.org/web/packages/EML/index.html>
- EML Python library: <https://pypi.org/project/eml-parser/>
- Geopackage format: [specifications](#) and tools ([rgdal](#), [ArcGIS Pro](#), [ArcGIS Desktop](#), [Python](#)). [Federal Geographic Data Committee/Content Standard for Digital Geospatial Metadata \(FFGC/CSDGM\)](#)

Example Data Products

- [Data Package](#)
- [Report](#)

File Formats

I&A Team's best practices for data management are to publish our data in open format. This list represents typical file formats for data we publish:

- Tabular data
 - CSV (comma-delimited text)
 - TXT (text)
 - XLSX (Office Open XML)
- GIS data
 - TIF (georeferenced TIFF file)
 - GPKG (OGC GeoPackage; modern replacement for shapefile)
 - GDB (geodatabase)

(Avoid proprietary formats such as .SHP)
- Text documentation
 - PDF
 - TXT
- Images
 - JPEG
 - PNG
 - TIFF

(Avoid proprietary formats such as .HEIC)

- Audio
 - MPEG-3 (MP3)
- Video
 - MPEG-4 (MP4)

For more information see [NARA Guidance](#)

Event/Occurrence/Relationship Diagram

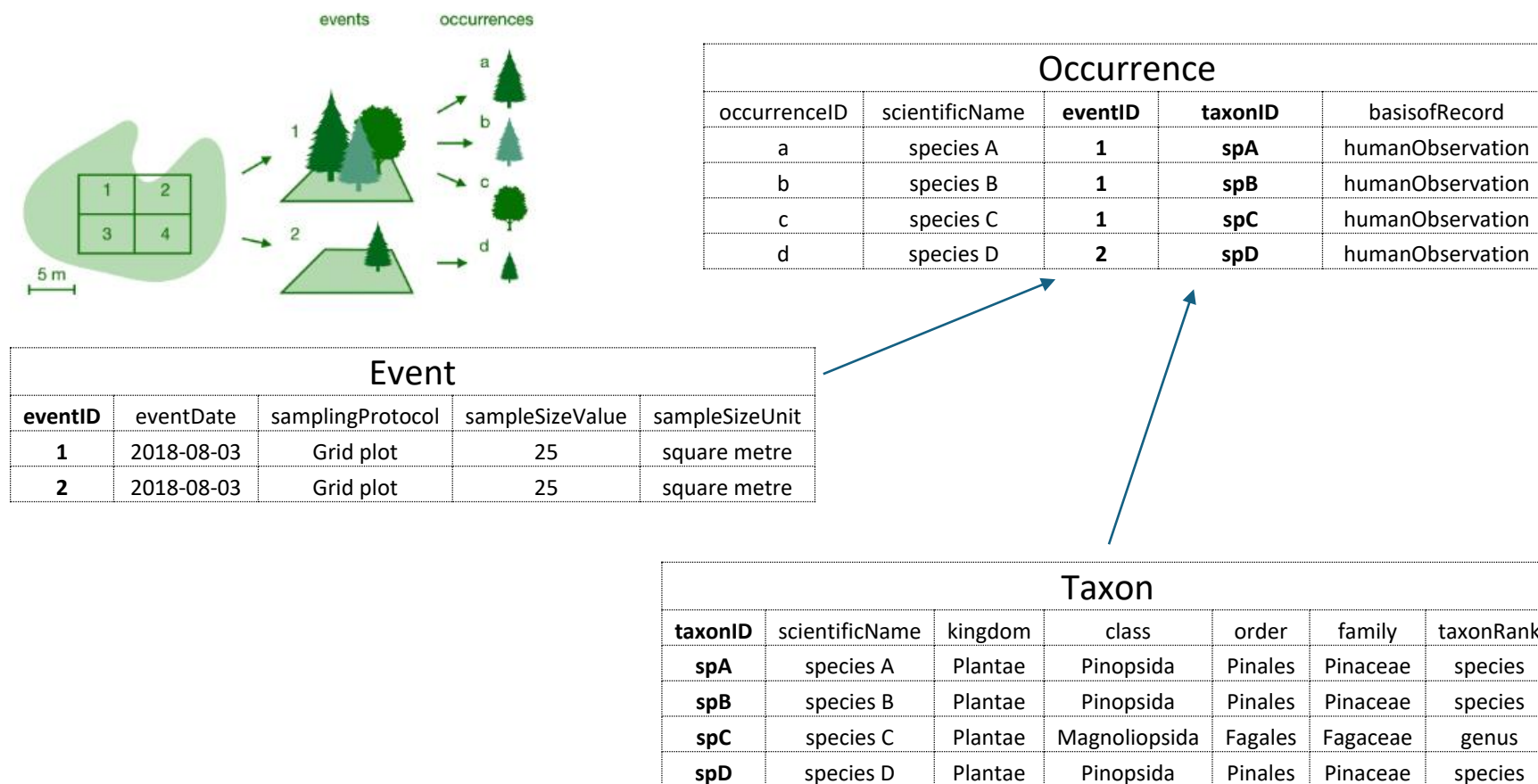


Diagram modified from [How to choose a dataset class on GBIF? - GBIF Data Blog](#)