




Documenting Spatial Data

Reina Chano Murray

Geospatial Data Curator & Applications Administrator
Johns Hopkins University

Sept 28, 2022



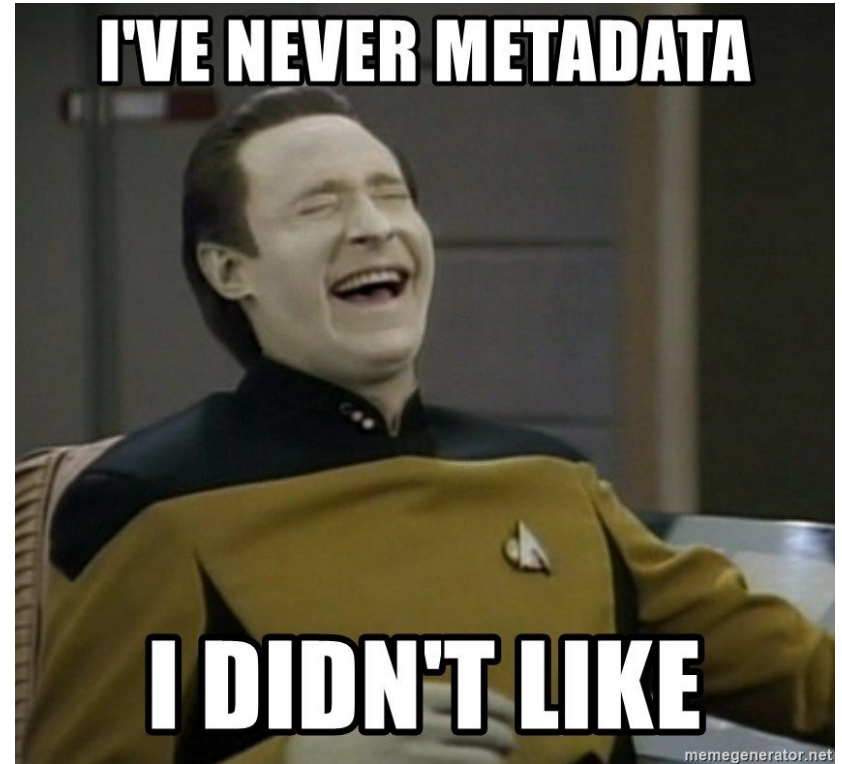
What We'll Cover

Metadata

- What is metadata?
- Why is it important?
- What are metadata standards?

Geospatial Metadata

- Geospatial metadata standards
- Geospatial metadata schemas
- Common elements to include
- Different levels of data
- Other recommendations for documentation (READMEs!)



Metadata

What is “metadata”?

*Metadata is **data** about **data**. It describes information like who collected it, when, for what purpose, and the level of quality... You can think of metadata as **little messengers to the future**.*

- Sarah Wakamiya, 2018

<https://www.nps.gov/articles/what-in-the-world-is-metadata.htm>



What is “metadata”?

- Provides context to data
 - making it easier to retrieve, use or manage the data in question
- Has some structure:

*“the metadata is collected so that it can fulfill a useful purpose, and sorted into known categories. It is this notion of structure that turns raw information into **actionable metadata**.”*

Jenn Riley, NISO, [Understanding Metadata: What is Metadata, and What is it For?: A Primer](#), 2017)

What is “metadata”?

Descriptive metadata	For finding or understanding a resource
Administrative metadata <ul style="list-style-type: none">- Technical metadata- Preservation metadata- Rights metadata	<ul style="list-style-type: none">- For decoding and rendering files- Long-term management of files- Intellectual property rights attached to content
Structural metadata	Relationships of parts of resources to one another
Markup languages	Integrates metadata and flags for other structural or semantic features within content

Jenn Riley, NISO, [Understanding Metadata: What is Metadata, and What is it For?: A Primer](#), 2017, page 6)

Why is metadata important?

Metadata makes your data more discoverable and understandable

What do we mean by “discoverable”?

[Feedback](#)

Geoportal

[About](#) [News](#) [Help](#)[Search](#)[Advanced](#)[Clear search](#)

You searched for:

Place > Philadelphia, Pennsylvania

[×](#)

Resource Class > Imagery

[×](#)

Resource Type > Raster data

[×](#)

Year > 1995 to 1998

[×](#)

Limit your search

1 - 2 of 2

Sort by Relevance

20 per page

[List view](#)[Map view](#)

Place

Pennsylvania

2

Philadelphia,

2

Pennsylvania

[×](#)

Resource Class

Imagery

[×](#)

2

Resource Type

Raster data

[×](#)

2

Year

1995 to 1998

2



Aerial Photography 1996 - Tile Index: Philadelphia, Pennsylvania

1.

1996. City of Philadelphia, Pennsylvania. Philadelphia aerial photography.



Aerial Photography 1996: Philadelphia, Pennsylvania

2.

1996. City of Philadelphia, Pennsylvania. Philadelphia aerial photography.

Why is metadata important?

Good metadata helps others

trust,

validate,

reuse

and

build upon

your data.

Why is metadata important?

*Tom Chatfield, a data management expert with the Bureau of Land Management bravely shared one real example from his own agency. A particular BLM field office developed an overall resource management plan for its region which included a map of oil and gas potential for its area. Naturally, this is an issue that registers strong opinions both for and against oil and gas recovery on public lands. When reviewed in the advent of litigation, BLM discovered that **no Data Quality information existed for the map**. No data sources could be identified, nor was there any discussion of the analysis methodology by which the oil and gas potential information was developed for the map. No geologic or geographic data were identified.*

*This was potentially disastrous in terms of defending the resource management plan. BLM was forced to develop a brand new map and closely document the provenance of the source data and methodology used to determine the potential ratings on the map. The **cost exceeded several thousands of dollars in staff time and production costs**, not to mention the **intangible costs of loss of trust in the agency by partners and the public**.*

Zolly, L.S., Henkel, H.S., Hutchison, V.B., Langseth, M.L., Thibodeaux, C.J., 2015, USGS data management training modules—metadata for research data [Slide 10]: U.S. Geological Survey, <https://doi.org/10.5066/F7RJ4GGJ>.



What are metadata standards?

- Provides agreement on the meaning and semantics of the data
- Think of metadata standards as a set of rules, or an agreement, that specify the minimum amount of information that should be documented about a dataset (and how)
- standardization makes it easier for **both humans** and **computers** to **find** and **understand** your data
- Metadata standards include **schemas**
 - schemas provide the overall structure for the metadata - provide a set of elements that should be used to describe a dataset

How many standards are there?

- So. Many.
- Check out Jenn Riley's Metadata Map: <http://jennriley.com/metadatamap/>
- Other sources
 - <https://rdamsc.bath.ac.uk/>
 - <https://fairsharing.org/>
 - <https://datastandards.directory/>
 - <https://www.iso.org/home.html>

Commonly Required Core Elements

- **Title** – The name of the resource. Titles should include place names and dates when available.
- **Creator** – The person(s) or organization(s) that created the resource.
- **Date** – The publication date for the resource.
- **Subject** – Theme or topic keywords of the resource chosen from a controlled vocabulary such as [Library of Congress Cartographic Genre](#) and/or [ISO Topic Category](#) terms.
- **Description** – A description of the resource that may include additional relevant information, like data creation methods, data sources, and special licenses. If the description is minimal or lacking, it can be improved by concatenating available metadata fields, such as title, date, format, and place.
- **Rights** – Values such as "Public" or "Restricted" if applicable, as well as any licensing information.



Geospatial Metadata



Common Geospatial Metadata Standards


ISO 191xx	"In 1999 the International Organization for Standardization (ISO) Technical Committee (TC) 211 Geographic Information / Geomatics was tasked to harmonize the FGDC Content Standard for Digital Geospatial Metadata (CSDGM) with other formal and defacto standards that support the documentation of geospatial data and services. The resultant ISO 19115: Geographic information - Metadata standard was finalized in 2003 and endorsed by the FGDC in 2010. A series of additional ISO 191** standards have been developed to update, extend, and supplement the 19115 standard."
FGDC CSDGM	"Content Standard for Digital Geospatial Metadata (CSDGM), Vers. 2 (FGDC-STD-001-1998) is the current version of this FGDC authored and endorsed standard. Executive Order 12906, directed that Federal agencies document geospatial resources using this standard. <i>Since the publication of EO 12906, the FGDC has endorsed several ISO Geospatial Metadata Standards that are now encouraged for use.</i> "
ArcGIS Metadata	"The ArcGIS metadata format contains elements that can store all content in all metadata standards that are supported by ArcGIS Desktop. It includes all Federal Geographic Data Committee (FGDC) <i>Content Standard for Digital Geospatial Metadata (CSDGM)</i> metadata content, all ISO 19139 metadata content, all North American Profile metadata content, and all INSPIRE metadata content."
ESRI-ISO	An older metadata standard from ArcGIS Desktop that can be upgraded to the ArcGIS metadata format
INSPIRE	The European Union standard for spatial data sets and services

Common Elements to Include

- Bounding box
- Geographic location
 - (place keywords -- such as <https://www.geonames.org/>)
- Spatial representation / geometry type
- Projection / coordinate reference system
- Provenance (how the data was made)
 - Geoprocessing workflow: geoprocessing tools used to create data
 - System/Software environment: software and software version used to create data
- Geospatial metadata schema used
- Maintenance schedule (if data will be updated)

NOTE: some of this get automatically taken care of by your GIS; others require manual entry

Example of a Geospatial Metadata Schema



Home

Repositories

Schema

Resources

Q Search OpenGeoMetadata

Open GeoMetadata (OGM) Aardvark Schema

[About OGM Aardvark](#) [How to Upgrade from GBL 1.0](#) [Legacy Schemas](#)

Fields by Group

Descriptive

- Title (R)
- Alternative Title
- Description (S)
- Language

Credits

- Creator (S)
- Publisher (S)
- Provider (S)

Categories

- Resource Class (R)
- Resource Type (S)
- Subject
- Theme
- Keyword

Temporal

- Temporal Coverage (S)
- Date Issued
- Index Year (S)
- Date Range

Spatial

- Spatial Coverage (S)
- Geometry (S)
- Bounding Box (S)
- Centroid
- Georeferenced

Relations

- Relation
- Member Of
- Is Part Of
- Source
- Is Version Of
- Replaces
- Is Replaced By

Rights

- Rights (S)
- Rights Holder
- License
- Access Rights (R)

Object

- Format (C)
- File Size

Links

- References (S)
- WxS Identifier

Identifiers

- ID (R)
- Identifier (S)

Admin

- Modified (R)
- Metadata Version (R)
- Suppressed

(R) = Required
(C) = Conditionally Required
(S) = Suggested

Fields in Alphabetical Order

Label	URI	Obligation
Access Rights	https://opengeometadata.org/docs/ogm-aardvark	Required

This site uses [Just the Docs](#), a documentation theme for Jekyll.

<https://opengeometadata.org/docs/ogm-aardvark>

Different Levels of Data to Document

- File-level
- Geodatabase/collection-level
- Presentation-level

Collection/Geodatabase-level + File-level

NRIS

NRIS_CR_Standards_Public.gdb

Historic_Buildings

Historic_Districts

Historic_Objects

Historic_Sites

Historic_Structures


Metadata

Geography

Table

National Register of Historic Places - National Geospatial Data Asset (NGDA) NPS National Register Dataset

Type <https://irma.nps.gov/DataStore/DownloadFile/647226>



Tags

NPS, National Park Service, NR, National Register of Historic Places, NRIS, CR, Cultural Resources

Summary

A current, accurate spatial representation of all historic properties listed on the National Register of Historic Places is of interest to Federal agencies, the National Park Service, State Historic and Tribal Historic Preservation Offices, local government and certified local governments, consultants, academia, and the interested public. This interest stems from the regulatory processes of managing cultural resources that are consistent with the National Historic Preservation Act as Amended (NHPA), the National Environmental Policy Act as Amended, the Archaeological Resources Protection Act, and other laws related to cultural resources. The regulations promulgating these laws require the use of spatial data in support of various decisions and actions related to cultural resource management.

Collectively, the National Register geo-spatial dataset is intended to be a comprehensive inventory of all cultural resources that are listed on the National Register of Historic Places. This dataset provides feature geometry representations (point, line, or polygon) and is intended to be supplemented with descriptive attributes maintained by other external database systems such as the National Register Information System which is included in this GeoDatabase.

Credits

Cultural Resources GIS, NPS
National Register of Historic Places, NPS

Use limitations

Properties listed on the National Register which are less than 10 acres are represented as points, while properties or districts which are more than 10 acres are delineated as polygons. All points and polygons are generated from coordinate pairs submitted with each individual National Register nomination. At this time, the quality and accuracy of the data is variable. Generally, listed historic properties represented as points (locations) are accurate to being placed in the appropriate county. Listed historic properties represented as polygons are far more problematic. Transcription errors and incorrect sequencing of bounding coordinates permeate this part of the dataset. Potential users are advised to carefully review the quality of the location and boundary of listed historic properties of interest to them before using this data. Subsequent releases of this dataset will seek to address these issues as time and funds are made available.

NRIS

NRIS_CR_Standards_Public.gdb

Historic_Buildings

Historic_Districts

Historic_Objects

Historic_Sites

Historic_Structures


Metadata

Geography

Table

Historic_Structures

Type File Geodatabase Feature Dataset



Tags

National Register of Historic Places, National Park Service, structures

Summary

The Historic Structures dataset contains the point and polygon representations of resources defined as structures in the National Register of Historic Places Inventory System (NRIS).

Description

Per the National Register Bulletin 16A, the term "structure" is used to distinguish from buildings those functional constructions made usually for purposes other than creating human shelter.

Examples: bridges, tunnels, gold dredges, firetowers, canals, turbines, dams, power plants, corncribs, silos, roadways, shot towers, windmills, grain elevators, kilns, mounds, cairns, palisade fortifications, earthworks, railroad grades, systems of roadways and paths, boats and ships, railroad locomotives and cars, telescopes, carousels, bandstands, gazebos, and aircraft.

Credits

There are no credits for this item.

Use limitations

There are no access and use limitations for this item.

Extent

West -175.134618 East 173.055623

North 65.349098 South -14.282311

Stutts M. 2014. National Register of Historic Places. National Register properties are located throughout the United States and their associated territories around the globe.

<https://irma.nps.gov/DataStore/Reference/Profile/2210280>

Presentation-level

HARVARD Geospatial Library


HARVARD LIBRARY

LoginBookmarksHistory

Try "population data" or "Cambridge"Clear All

Historic County Boundaries and Total Population, United States, 1830

WEB MAP SERVICE (WMS)



Click map to inspect values

Leaflet | OpenStreetMap contributors, Carto

AUTHOR(S) University of Minnesota, Minnesota Population Center

DESCRIPTION This boundary file contains historic county boundaries for which the U.S. Census Bureau tabulated data and was produced by the Minnesota Population Center as part of the National Historical Geographic Information System (NHGIS) project. The 1830s county boundaries depicted were paired with 1830s total population data from the same year by the Harvard University Center for Geographic Analysis. The county boundary datasets and its respective population data were paired for each decade, 1790-2000. The 1830s is the National Historical Geographic Information System (NHGIS) to create a digital spatial temporal database of all available historical U.S. geographic census materials. The available datasets on the

MORE DESCRIPTION

PUBLISHER Minnesota Population Center

PLACE(S) United States

SUBJECT(S) Boundaries
Census
Census
Population
Boundaries
History

MORE SUBJECTS

YEAR 1830

HELD BY Harvard

TOOLS & RELATED LINKS

DOWNLOAD SHAPFILE

Click
Mapview
Open in Carto
Bookmark

ABOUT THIS ITEM

Public
Harvard
Public

HARVARD Geospatial Library

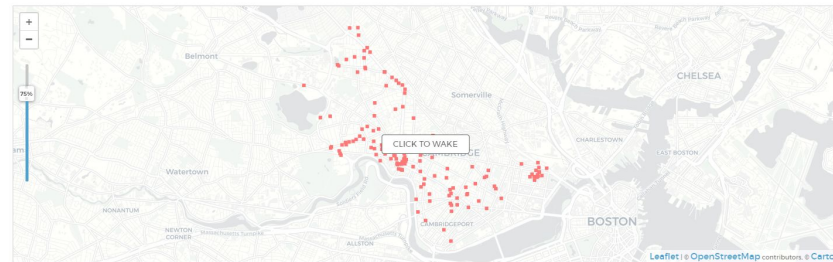
HARVARD LIBRARY

LoginBookmarksHistory

Try "population data" or "Cambridge"Clear All

Historical Markers, Cambridge, Massachusetts, 2014

WEB MAP SERVICE (WMS)



AUTHOR(S)


Cambridge (Mass.) Geographic Information Systems

TOOLS & RELATED LINKS

<https://hgl.harvard.edu/catalog/harvard-nhgis-pop1830>

Presentation-level

Active Hurricanes, Cyclones and Typhoons



This layer describes the observed path, forecast track, and intensity of tropical cyclone activity (hurricanes, typhoons, cyclones) from the National Hurricane Center (NHC) and Joint Typhoon Warning Center (JTWC).

Manage by: [arc/landinfo](#)

Created: Jun 11, 2019 Updated: Sep 28, 2022 View Counts: 92,214,393

[Subscribe](#) [Living Atlas](#)

Overview Data Visualization

Open in Map Viewer Classic

Open in Scene Viewer

Open in ArcGIS Desktop >

Share

Metadata

Description

Hurricane tracks and positions provide information on where the storm has been, where it is currently located, and where it is predicted to go. Each storm location is displayed by the National Hurricane Center (NHC) and the Joint Typhoon Warning Center (JTWC). Only applies to hurricanes in the Atlantic and Eastern Pacific basins, however all storms are still symbolized using that classification for consistency.

Data Source

- This data is provided by NOAA National Hurricane Center (NHC) for the Central-east Pacific and Atlantic, and the Joint Typhoon Warning Center for the West-Central Pacific and Indian basins. For more information, see [NHC's Data Users Guide](#).

Sample Data

- See [Sample Layer Item](#) for sample data during inactive Hurricane Season.

Update Frequency

- The [Aggregated Line Reads](#) methodology checks the Source for updates every 15 minutes. Tropical cyclones are normally updated every 6 hours at 5:00 AM EDT, 11:00 AM EDT, 5:00 PM EDT, and 11:00 PM EDT (or 4:00 AM EST, 10:00 AM EST, 4:00 PM EST, and 10:00 PM EST).
- Public advisories for Eastern Pacific tropical cyclones are normally updated every 6 hours at 2:00 AM PST, 8:00 AM PST, 2:00 PM PST, and 8:00 PM PST (or 1:00 AM PST, 7:00 AM PST, 1:00 PM PST, and 7:00 PM PST).
- Intermediate public advisories may be issued every 2 hours when coastal watches or warnings are in effect, and every 2 hours when coastal watches or warnings are in effect and land-based radar has identified a suitable storm center. Additionally, special advisories may be issued at any time due to significant changes in warnings or a cyclone. For the NHC, data source you can subscribe to RSS Feeds.
- North Pacific and North Indian Ocean tropical cyclone warnings are updated every 6 hours, and South Indian and South Pacific Ocean tropical cyclone warnings are normally updated every 12 hours. These are with 24-hour.

Scale/Resolution

- The horizontal accuracy of these datasets is not stated but it is important to remember that tropical cyclone track forecasts are subject to error, and that the effects of a tropical cyclone can span many hundreds of miles from the center.

Area Covered

- World

Glossary

- Forecast location:** Represents the official NHC forecast locations for the center of a tropical cyclone. Forecast center positions are given for projections and 12, 24, 36, 48, 72, 96, and 120 hours after the forecast nominal initial time. Click here for more information. Forecast points from the JTWC are valid 12, 24, 36, 48, and 72 hours after the forecast initial time.
- Forecast track:** This product aids in the visualization of an NHC official track forecast, the forecast points are connected by a red line. The track line is not a forecast product, as such, the line should not be interpreted as representing a specific forecast for the location of a tropical cyclone in between official forecast points. It is also important to remember that tropical cyclone track forecasts are subject to error, and the effects of a tropical cyclone can span many hundreds of miles from the center. Click here for more information.
- The Cone of Uncertainty:** Cyclone paths are hard to predict with absolute certainty, especially days in advance. The cone represents the probable path of the center of a tropical cyclone and is formed by enclosing the area swept by a set of circles along the forecast track (at 12, 24, 36, 48, and 72 hours, etc). The size of each circle is scaled so that two-thirds of the historical official forecast errors over a 5-year sample fall within the circle. Based on the previous 5 years, the error track of a tropical cyclone can be expected to remain within the cone roughly 60-70% of the time. It is important to note that the area affected by a tropical cyclone can extend well beyond the confines of the cone enclosing the most likely track area of the center. Click here for more information.
- Coastal Warning:** Coastal areas are placed under watches and warnings depending on the proximity and intensity of the approaching storm.
 - Tropical Storm Watch is issued when a tropical cyclone containing winds of 34 to 63 knots (39 to 73 mph) or higher poses a possible threat, generally within 48 hours. These winds may be accompanied by storm surge, coastal flooding, and/or rain flooding. The watch does not mean that tropical storm conditions will occur, it only means that these conditions are possible.
 - Tropical Storm Warning is issued when sustained winds of 34 to 63 knots (39 to 73 mph) or higher associated with a tropical cyclone are expected to 24 hours or less. These winds may be accompanied by storm surge, coastal flooding, and/or rain flooding.
 - Hurricane Watch is issued when a tropical cyclone containing winds of 64 knots (74 mph) or higher poses a possible threat, generally within 48 hours. These winds may be accompanied by storm surge, coastal flooding, and/or rain flooding. The watch does not mean that hurricane conditions will occur, it only means that these conditions are possible.
 - Hurricane Warning is issued when sustained winds of 64 knots (74 mph) or higher associated with a tropical cyclone are expected to 24 hours or less. These winds may be accompanied by storm surge, coastal flooding, and/or rain flooding. A hurricane warning can remain in effect after progressively high water or a combination of dangerously high waves and exceptionally high waves continue, even though winds may be less than hurricane force.

Revisions

- Mar 27, 2022: Added US, Mac, SE, Mac, Wind, Mac, Gov, and Mac, Label, Label to Forecast and Cone layers.

This map is provided for informational purposes and is not monitored 24/7 for accuracy and currency. Always refer to NOAA or JTWC sources for official guidance.

If you would like to be alerted to potential issues or simply use when this Service will update next, please visit our [Live Feed Status Page](#)

Active Hurricanes, Cyclones and Typhoons

ISO-19139 Metadata

Metadata Information:

Metadata language: en

Metadata character set: utf8

Last update: 2022-08-16

Metadata contact - pointOfContact:

Organization's name: Eri

Scope of the data described by the metadata: dataset

Scope name: dataset

Name of the metadata standard used: ISO 19139 Geographic Information - Metadata - Implementation Specification

Version of the metadata standard: 2007

Metadata identifier: 248e7b5827a34b248647afb012c58787

Resource Identification Information:

Citation:

Title: Active Hurricanes, Cyclones and Typhoons

Reference date - publication: 2022-08-16

Themes or categories of the resource: climatology/Meteorology/Atmosphere, environment

Place keywords:

Keywords: United States of America

Abstract:

Hurricane tracks and positions provide information on where the storm has been, where it is currently located, and where it is predicted to go. Each storm forecast points from the NHC are valid 12, 24, 36, 48, and 72 hours after the forecast's initial time. Forecast track: This product aids in the visualization of the cone represents the probable track of the center of a tropical cyclone and is formed by enclosing the area swept out by a set of circles along the forecast track.

Purposes:

This layer describes the observed path, forecast track, and intensity of tropical cyclone activity (hurricanes, typhoons, cyclones) from the National Hurricane Center (NHC) and the Joint Typhoon Warning Center (JTWC).

Dataset language: en

Dataset character set: utf8

Status: completed

Resource constraints:

Constraints:

Limitations of use: This work is licensed under the Eri Master License Agreement. View Summary | View Terms of Use

Resource constraints:

Legal constraints:

Access constraints: otherRestrictions.

Limitations of use: Other Constraints

Resource constraints:

Legal constraints:

Use constraints:

Limitations of use: Other Constraints

<https://www.arcgis.com/home/item.html?id=248e7b5827a34b248647afb012c58787>

Presentation-level

Search or jump to...

Pull requests Issues Marketplace Explore

datadesk / california-coronavirus-data Public

Watch 21 Fork 138

<> Code Issues 3 Pull requests 5 Actions Security Insights

master 6 branches 0 tags

Go to file Add file Code

actions-user Automatic update ✓ 469c92 1 hour ago 4,704 commits

.github/workflows	Update notebooks.yml	2 years ago
binder	Add altair-latimes dependency	3 years ago
notebooks	Flag the top growing counties	2 years ago
pdf/adult-and-senior-care	Automatic update	13 months ago
.gitignore	Better notebook tests and runners	2 years ago
LICENSE	Hello world	3 years ago
Pipfile	Updated Pipfile	2 years ago
Pipfile.lock	Updated lock file	13 months ago
README.md	Update README.md	10 months ago
cdc-vaccination-state-totals.csv	Automatic update	12 days ago
cdcr-prison-totals.csv	Automatic update	21 hours ago
cdcr-state-totals.csv	Automatic update	21 hours ago
cdph-adult-and-senior-care-facilities...	Automatic update	12 months ago
cdph-adult-and-senior-care-totals.csv	Automatic update	14 months ago
cdph-age.csv	Automatic update	5 days ago
cdph-county-cases-deaths.csv	Automatic update	5 days ago

About

The Los Angeles Times' open-source archive of California coronavirus data

www.latimes.com/coronaviustracker

python data news csv binder jupyter journalism pandas data-journalism altair coronavirus covid

Readme View license 208 stars 21 watching 138 forks

Contributors 17

+ 6 contributors

Languages

<https://github.com/datadesk/california-coronavirus-data>

Presentation-level

☰ README.md

The Los Angeles Times' open-source archive of California coronavirus data





 launch  binder  Jupyter Notebook tests  passing

Table of contents

Cases and deaths

- [cdph-county-cases-deaths.csv](#)
- [cdph-state-cases-deaths.csv](#)
- [cdph-positive-test-rate.csv](#)
- [cdph-variants](#)
- [latimes-agency-totals.csv](#)
- [latimes-county-totals.csv](#)
- [latimes-state-totals.csv](#)
- [latimes-place-totals.csv](#)
- [latimes-county-recovery-estimates.csv](#)
- [latimes-state-recovery-estimates.csv](#)

Vaccinations

- [cdc-vaccination-state-totals.csv](#)
- [cdph-vaccination-county-totals.csv](#)
- [cdph-vaccination-state-totals.csv](#)
- [cdph-vaccination-zipcode-totals.csv](#)

Demographics

- [cdph-age.csv](#)
- [cdph-population-race-ethnicity.csv](#)
- [cdph-race-ethnicity.csv](#)
- [cdph-vaccination-county-by-age.csv](#)
- [cdph-vaccination-county-by-race-ethnicity.csv](#)
- [cdph-vaccination-state-by-age.csv](#)
- [cdph-vaccination-state-by-race-ethnicity.csv](#)

<https://github.com/datadesk/california-coronavirus-data>

Presentation-level



GIS rasters to identify sites for creating habitat for American Woodcock in Rhode Island

Buffum, Bill, University of Rhode Island, <https://orcid.org/0000-0002-4312-0876>

buffum@uri.edu

Publication date: August 13, 2021

Publisher: Dryad

<https://doi.org/10.5061/dryad.pg4f4qrp6>

Citation

Buffum, Bill (2021), GIS rasters to identify sites for creating habitat for American Woodcock in Rhode Island, Dryad, Dataset, <https://doi.org/10.5061/dryad.pg4f4qrp6>

Abstract

The University of Rhode Island has conducted several studies of habitat use of *Scolopax minor* (American Woodcock) in Rhode Island, USA. In 2020 we developed a new species distribution model (SDM) tool to identify sites in the Rhode Island where forest clearcutting to create young forest habitat would have the most positive effect for American woodcock. A typical SDM predicts the probability of presence (POP) of a species at any location based on an analysis of known occurrences and environmental variables, but it cannot predict how much the POP of a species would change after a new patch of young forest is created in any location. We believe that our new tool is effective, and that it will help landowners identify the best locations on their properties to improve woodcock habitat. We also believe that similar tools can be developed for other wildlife species of conservation concern. We created the new tool by modifying the existing 2018 SDM raster for American Woodcock in Rhode Island. Creating the tool involved creating four new ArcGIS raster datasets. The existing 2018 SDM raster and the four new rasters are now publicly available in a geodatabase in the Dryad repository.

Methods

We used Maxent 3.4.1 to create our new SDM tool, and ArcGIS Desktop 10.6 to store the output raster datasets in a geodatabase that can only opened with geographic information system (GIS) software. The geodatabase uploaded to Dryad (Woodcock_POP_2018_Rasters) contains five raster datasets that cover the entire state of Rhode Island. A brief

Search

Explore Data | About ▼ | Help ▼ | Login

Data Files

Download dataset

June 15, 2021

README.txt	3.83 kB
Woodcock_PO...asters.zip	134.03 MB

Metrics

79 views

4 downloads

0 citations

Keywords

Other natural sciences

License

This work is licensed under a CC0 1.0 Universal (CC0 1.0) Public Domain Dedication license.

Methods

We used Maxent 3.4.1 to create our new SDM tool, and ArcGIS Desktop 10.6 to store the output raster datasets in a geodatabase that can only opened with geographic information system (GIS) software. The geodatabase uploaded to Dryad (Woodcock_POP_2018_Rasters) contains five raster datasets that cover the entire state of Rhode Island. A brief description of the five rasters is provided below. We provide a detailed description of how we created our new SDM tool, and how the tool can be used, in the following journal article: [Bill Buffum, Roger Masse and Scott R. McWilliams \(2021\) Novel Use of Species Distribution Modeling to Identify High Priority Sites for American Woodcock Habitat Management. Northeastern Naturalist 28: p. 233-247.](#)

Woodcock_POP_2018: This raster (10m pixels) shows the 2018 probability of presence for woodcock in RI. It is based on Maxent modelling of three years woodcock location data, obtained by catching 68 woodcock in RI in the spring, attaching transmitters, and tracking their movements until August. The Maxent model was projected onto 2018 spatial data.

Woodcock Upland POP_2018: This raster (50m pixels) shows the 2018 probability of presence for woodcock in RI. It is similar to Woodcock_POP_2018, but it has 50m pixels and only includes upland forest areas which are suitable for clearcuts to create young forest habitat.

Woodcock POP After Forest Mgt_2018: This raster (50m pixels) was the primary product of our new SDM tool. It shows the predicted probability of presence for woodcock in RI if young forest habitat is created at any upland forest location in RI.

Increase in Woodcock POP After Forest Mgt_2018: This raster (50m pixels) shows the predicted increase in probability of presence for woodcock in RI if young forest habitat is created at any site. It was produced by using raster calculator to subtract the POP values of Woodcock Upland POP_2018 from Woodcock POP After Forest Mgt_2018

Top Priority Sites for Forest Mgt: This raster (50m pixels) shows the top priority (top 20% and top 40%) sites for creating young forest habitat for woodcock in Rhode Island. The top 20% sites include area where both the predicted POP increase after forest management and the predicted POP after forest management are in the top 20% class. The top 40% sites include areas where both the predicted POP increase after forest management and the predicted POP after forest management are in the top 40% range.

Usage Notes

The ArcGIS raster datasets can only be viewed by using geographic information system (GIS) software such as ArcGIS Desktop or ArcGIS Pro (Environmental Systems Research Institute, Redlands, CA). The geodatabase, a compressed version of which is uploaded to Dryad, contains a total of 619 files, but when it is opened with GIS software, the five raster datasets will be visible and easily opened.

This work is licensed under a CC0 1.0 Universal (CC0 1.0) Public Domain Dedication license.



<https://datadryad.org/stash/dataset/doi:10.5061%2Fdryad.pg4f4qrp6>
<https://orcid.org/0000-0002-4312-0876>

Tips for Documentation

- READMEs are your best friend
- Document your data along the way - saves you time at the end
- Use descriptive file names
- If you're using geospatial desktop software or web GIS, create your metadata in the platform/software you start in (inheritance)
 - Convert metadata into XML

README Files - What is it?

A README file is a text file containing key information about your data which gives the reader a general understanding of the purpose and history of your data set, how it is organized, and how it can be used. You can think of a README file as a **manual for your data**.

You can use it to capture components of your data that are not adequately captured in the metadata contained with your geoprocessing application.

README Files - Why create one?

While some of the information contained within your README file may overlap with the content you entered in the metadata within your geoprocessing tool, it is still a good idea to create a separate file that lives outside of your geospatial file.

This is helpful for not only if you share your data with others, but is also beneficial to yourself if you need to revisit your data in the future.

Furthermore, some geospatial formats (ie, GeoJSON) or other formats you might save your file as (csv) will not store metadata like a shapefile.

Cover all your bases.

README Files - What to include?

- High-level/core information such as purpose of the data set, where the files can be found
- Geospatial-specific elements such as coordinate ref. system,
- Workflow/software environment such as software version, data manipulations which occurred outside of your geoprocessing tool, data version history
- Details about your files such as the purpose of key files and your file naming scheme

See <https://data.research.cornell.edu/content/readme>

OGC Implementation Standards



[ABOUT](#) [MEMBERSHIP](#) [STANDARDS & RESOURCES](#) [INNOVATION](#) [NEWS & EVENTS](#)

OGC Standards

Below is a list of OGC Implementation Standards.

Implementation Standards are different from the Abstract Specification. They are written for a more technical audience and detail the interface structure between software components. An interface specification is considered to be at the implementation level of detail if, when implemented by two different software engineers in ignorance of each other, the resulting components plug and play with each other at that interface.



Any Schemas (xsd, xslt, etc) that support an approved Implementation Standard can be found in the official [OGC Schema Repository](#).

[Standards List](#)

[Standards Architecture Diagram](#)

Standards Listing

A list of the OGC standards

Document Title (click to view/download)	Version	OGC Doc.#	Editor	Publication Date
 CDB Multi-Spectral Imagery Extension CDB: Multispectral	1.0	17-080r2	Ryan Franz	2018
The "Multi-Spectral Imagery" extension defines how to encode and store reflected electromagnetic radiation from the infrared wavelengths into a CDB. The portion of the spectrum targeted is between the visible spectrum (current imagery and texture in CDB), and longer wavelength infrared that is See more...				
 CF-netCDF3 Data Model Extension standard netcdf-data-model-ex	3.1	11-165r2	Ben Domenico and Stefano Nativi	2013

The OGC netCDF encoding supports electronic encoding of geospatial data, that is, digital geospatial information

OGC Standards

- [3D Tiles](#)
- [3dP](#)
- [ARML2.0](#)
- [Cat: eBRIM App Profile: Earth Observation Products](#)
- [Catalogue Service](#)
- [CDB](#)
- [CityGML](#)
- [CityJSON](#)
- [Coordinate Transformation](#)
- [EO-GeoJSON](#)
- [Filter Encoding](#)
- [GML in JPEG 2000](#)
- [GeoAPI](#)
- [GeoPackage](#)
- [GeoSciML](#)
- [GeoSPARQL](#)
- [Geography Markup Language](#)
- [GeoRSS](#)
- [Geospatial eXtensible Access Control Markup Language \(GeoXACML\)](#)
- [Geospatial User Feedback \(GUF\)](#)
- [GeoTiff](#)
- [GroundwaterML](#)
- [HDFS](#)
- [I3S](#)
- [Indoor Mapping Data Format \(IMDF\)](#)

Data Principles to Know

FAIR

<https://www.go-fair.org/fair-principles/>

CARE

<https://www.gida-global.org/care>

Open Data Charter

<https://opendatacharter.net/principles/>

OVERLAP BETWEEN			
FAIR and CARE Principles			Open Data Charter Principles
FAIR	<u>F</u> indable		Timely and Comprehensive
	<u>A</u> ccessible		Accessible and Usable
CARE	<u>I</u> nteroperable		Comparable and Interoperable
	<u>R</u> eusable		
	<u>C</u> ollective Benefit		For Improved Governance & Citizen Engagement
	<u>R</u> esponsibility		For Inclusive Development and Innovation
	<u>A</u> uthority to Control		
	<u>E</u> thics		Open By Default



Questions?



Reina Chano Murray

<https://github.com/reinacmurray>

reinamurray@jhu.edu