



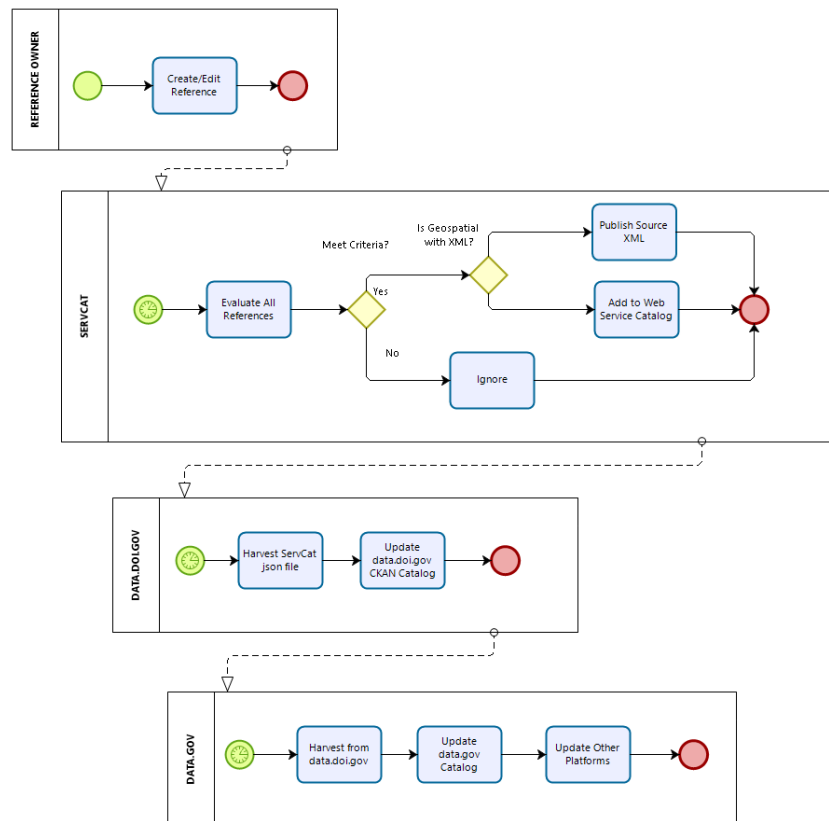
Publishing ServCat Content to data.gov

Last Saved: March 17, 2020

As part of [Project Open Data](#), [ServCat](#) publishes a subset of its entire catalog to [data.gov](#)¹, where it is discoverable to other communities that are not aware of ServCat. This document describes the process and criteria for publishing information to this clearinghouse for the purpose of making our information even more accessible.

1 Process

Publishing content to data.gov actually involves four different steps, each of which is dependent on, but asynchronous with the other.



- **Reference Owners** - ServCat content is added and updating regularly.
- **ServCat** - Once a day, ServCat evaluates all content it manages to see if it meets the criteria (Section 2 below) to publish. Content meeting the criteria and not a geospatial record created

¹ Only metadata is published to data.gov – the files still remain in ServCat

by XML is published as a text file, conforming to version [1.1](#) of the project open data common core metadata standard https://ecos.fws.gov/ServCat/OpenData/FWS_ServCat_v1_1.json². Geospatial records created by XML and also meeting the criteria are published separately to <https://ecos.fws.gov/ServCatGos/files/> for harvesting by the [geospatial one stop](#).

- data.doi.gov – Once a day data.doi.gov harvests from ServCat and synchronizes its catalog. This harvest is orchestrated by the USFWS Branch of Geospatial Data Services who help (1) defining the criteria of what should be harvested to data.doi.gov, and (2) configure data.doi.gov to harvest from the ServCat endpoint.
- data.gov – The harvest from data.doi.gov has been periodic due to changes with contracting. This should resume to a daily harvest in the near future.

2 Criteria

A ServCat record must meet all of these criteria:

- **Active Lifecycle** – Draft and Inactive references are ignored
- **Public File Access** – Associated files must be downloadable by the public
- **Valid Reference Type** – Only datasets (either spatial or tabular) and maps are valid
- **Created for/by the USFWS** – The USFWS must be the originator of the information
- **High Quality** – The data must be of high quality
- **Has at Least One Linked Resource that is not a PDF** – This resource can be a digital file, web service, or external link but may not be a PDF
- **Has Date of Issue** – The date of issue may not be empty. If the reference is a project, the project start/end dates must have at least one date value.
- **Not Older Version** – References marked as older versions are excluded
- **Has Description** – The description may not be empty
- **The Title is not empty, 'Untitled' or 'No Title'** – This goes without saying...

² This url will only work on computers that are **off** the FWS network.

3 Tools

In case you are interested, the following report summarizes what references are made available for harvesting to data.doi.gov:

<https://ecos.fws.gov/ServCat/DataStoreReports/Contributors/Project%20Open%20Data>

The current FWS content available in data.doi.gov is found here:

<https://data.doi.gov/dataset?organization=fws-gov>

The current FWS content available in data.gov is found here:

https://catalog.data.gov/dataset?q=&sort=score+desc%2C+name+asc&publisher=Fish+and+Wildlife+Service&organization_type=Federal+Government&organization_limit=0&organization=doi-gov

4 Field Crosswalk

The following table summarize the crosswalk between the Open Data v1.1 standard and ServCat.

| Open Data v1.1 | ServCat Field | Notes/Logic |
|----------------|---|---|
| type | 'Dataset' | Hard-coded |
| title | Title | |
| description | Abstract/Description | |
| keyword | Keyword(s), Subject Categories, Taxonomy Common Name(s), Taxonomy Scientific Name(s) | If nothing from above items, then Title is parsed into keywords with articles (the, and, etc.) removed |
| modified | Date Of Issue, Content Begin Date, Content End Date | Date of Issue, else Content Begin, else Content End, formatted as YYYY-MM-DD Current criteria guarantee at least one of those will have a value. |
| publisher | 'Fish and Wildlife Service' | Hard-coded |
| contactPoint | 'Brent Frakes' | Hard-coded |
| Identifier | 'FWS_ServCat' [ReferenceCode] | |
| accessLevel | 'Restricted Public' 'Public' | IF File Access Level != Public THEN ' <i>restricted public</i> ' ELSE 'public' |
| programCode | 010:094, 010:028 | Hard-coded |
| bureauCode | 010:18 | Hard-coded |
| rights | Access Level | IF ' restricted public ' THEN ' <i>The information resource is proprietary and/or copyrighted. Access to this information by the public is very limited. Please contact the associated</i> |

| | | |
|---|---|--|
| | | <p><i>Point of Contact regarding additional use constraints and requirements for obtaining the information.</i></p> <p><i>ELSE <Blank></i></p> |
| spatial | Bounding Boxes, Units | <p>If the Reference is linked to one or more geospatial bounding boxes, union all of the bounding boxes into a single geospatial Envelope.</p> <p>If there are no bounding boxes, take all the Unit codes (current criteria guarantee at least one), retrieve their bounding boxes from the Unit REST service, and union them all into a single geospatial Envelope.</p> <p>If we have a geospatial Envelope, it is conveyed as two sets of coordinates (four total coordinates) representing the Bottom Left and Top Right (Min Long, Min Lat, Max Long, Max Lat).</p> <p>If we do not have an Envelope, then use the Unit REST Service to supply the full name fo the first Unit. The idea is that this serves as a geotag that the harvest can translate from http://www.geonames.org/</p> |
| temporal | Content Begin Date, Content End Date | A date range supplied only if both Content Begin and Content End are available. Formatted as YYYY-MM-DD/YYYY-MM-DD |
| dataQuality | True False | A 'True' or 'False' value derived from the 'Quality' assigned to the Reference. If Quality is explicitly tagged as 'High', this field is 'true'. False otherwise. |
| issued | Date Of Issue | A date supplied only if Date of Issue is available. Formatted as YYYY-MM-DD. |
| landingPage | Reference Profile URL | |
| references | External Link URL | This is a json collection of strings. The url's of external websites linked to the Reference, if any. And always the url of the Reference Profile Page. |
| theme | Reference Type | <p>"This is a json collection of strings. However, we only provide one string. It is the 'readable' (display name) of the Reference Type.</p> <p>Note: Per the Geoplatform documentation, if this collection contains an entry with the single word 'Geospatial', the record will be pulled by Geoplatform. While we do have Reference Types with the word 'Geospatial' in them, we do not have a Reference Type made up of only that word. Likely, therefore, Geoplatform is not pulling any of the appropriate Geospatial dataset records."</p> |
| distribution | External Links; Digital Files, Web Services | This is a json collection of one or more items describing a download source. Each item has multiple properties, as described in the next section. |
| One distribution record per each digital file linked to reference | | |
| type | | Hard-coded |
| downloadURL | File Download URL | |
| mediaType | File Type | A mime-type string. We supply these from the extension of the file name (e.g. 'pdf', 'zip', 'txt'). If the file has no extension, a default value is supplied. |

| | | |
|--|-------------------------|---|
| format | File Type | This is just the file extension itself (e.g. 'pdf', 'zip', 'txt') |
| description | File Description | Any description supplied for the file |
| title | File Name | |
| One distribution record per each web service linked to reference | | |
| type | | Hard-Coded |
| accessURL | Web Service URL | |
| mediaType | <blank> | Hard-coded |
| format | 'API' | Hard-coded |
| description | Web service description | |
| title | Web Service Type | Type of web service (e.g. Map Service, Image Service, Data Service) |