



# HDX

# Humanitarian Data Exchange

## Providing metadata for your CODs\* on HDX

\* [IASC Common Operational Datasets](#)

### Introduction

A popular definition of metadata is data about data or data that defines or describes other data. Metadata can be thought of as additional information or documentation about your dataset that will make it easy for others to understand and to put your data into context. Datasets on HDX must provide a minimum set of metadata that you must fill in for each dataset you share. The following table lists these metadata fields and provides guidance on how to fill them in. The metadata fields in **red bold italic** are mandatory and must be filled in, while the rest are optional. The instructions in **blue** are specific to CODs and FODs and have been added to the original HDX metadata guidance.

### General tips when adding COD FOD metadata to HDX

- The search algorithm will look for any match in the following metadata elements: Title, Description, Tags
- The more information you can put in the metadata the more operational data is understood and used properly. Indicating challenges and the purpose of data really helps and may even provide opportunities for improvement and collaboration
- CODs are the “best available” datasets that are agreed to be used in a humanitarian response. It is better to share data that is incomplete and let partners know about it than to not share it.

### Dataset metadata fields

Metadata field	Guideline for using the field
<b>Location</b>	Use this field to list the location(s) to which your data relates. You can select one or more locations. If your data is more or less global in nature, you may select ‘World’ as the location.  Example: Uganda
<b>Dataset Title</b> <b>This metadata element is used in the search algorithm.</b>	Use this field to provide short title that concisely describes your dataset. The title of the dataset is listed in search results and in order to help other users find the right dataset, it is important that you provide key clues about the location, subject matter and time period of your dataset in the title. Avoid using context specific abbreviations in the title which may not be known outside of your context. Also avoid using words that are relative such as current, latest or previous when referring to the time period (e.g Latest 3W) because such terms become misleading as the dataset ages. Example: Who is Doing What Where in Afghanistan, Jan Dec 2014 COD <ul style="list-style-type: none"><li>• Include the term COD in the title so clients are aware this is the authoritative datasets used in operations.</li></ul> Example: “Nepal admin level 1 administrative boundaries (Sep 2015)”; “Mali Health Facilities (2013)”
<b>Source</b>	Use the source to attribute the source of the data. The source of the data is the entity responsible for <i>collecting or generating</i> the data. The source should not be confused with the “organization” (see below) which is the organization on HDX responsible for <i>sharing</i> the data on HDX. Use acronyms rather than full names of the source where possible. In those cases where the data is sourced from more than one source, list all the sources separated by commas. You may enter “Multiple Sources” in the field if the list of sources is too long in which case, make sure you attribute the sources in an additional metadata file. Example: FAO, UNICEF, WHO COD <ul style="list-style-type: none"><li>• Do not use local level acronyms, write these out so people outside or new to the context can understand, provide a weblink</li></ul>

	<p>Example: "Ukrstat" should be "State Statistics Service of Ukraine"</p>
<p><b>Description</b> This metadata element is used in the search algorithm.</p>	<p>Use this field to provide a more detailed description of the dataset. The description should clarify the three aspects (location, subject matter and time period) mentioned in the title. For instance, you may specify the subnational coverage of a dataset or list the indicators that are contained in the dataset. Mention how (if at all) the dataset will be updated and consider providing a link to where users can expect to find the most updated version of the data. If you make use of an additional metadata file, then mention this fact in the description field. You may also provide links to the source website of the data if one is available. Avoid long descriptions and aim for no more than 2 paragraphs. If the field names (i.e. column headers, attributes) in your dataset are not easy to interpret, you can define them here or in a separate metadata file that you upload.</p> <p>Example: This dataset contains Who, What, Where (3W) data for Afghanistan. The operational presence of the various organisations (who) by sector (what) and location (where) at district level is presented. The dataset covers activities carried out in 2014. The 3W for Afghanistan is updated on an annual basis, and the most up-to-date version of the data can be found at <a href="http://www.example.url">www.example.url</a> . Please refer to the readme.txt file for complete metadata.</p> <p><b>COD</b></p> <ul style="list-style-type: none"> <li>• Include information about the following: spatial details (projection, scale, polygon, line, point, raster, etc); description of items if they are not clear; purpose/ advantages to the dataset (e.g. very accurate for a particular reason); For administrative boundaries include the name of the admin level (e.g., Region, Nahya, Barangay, Cercles), the data sponsor.</li> <li>• Disaster specific data (e.g. earthquake epicentre) Include <a href="#">GLIDE number</a></li> </ul> <p>Example: This dataset depicts the Health Infrastructure of Nepal as points the scale is 1/10,000.. The source of the data Survey Department of Nepal (<a href="http://ngiip.gov.np/">http://ngiip.gov.np/</a>) and data sponsor is WHO (World Health Organization). This dataset is missing locations in the north of the country.</p>
<p><b>Dataset contains subnational data</b></p>	<p>Use this field to let other users know whether or not your dataset contains data that is disaggregated beyond the national level. The field is checked by default. Remember to uncheck it if your data is not disaggregated below the national level</p>
<p><b>License</b></p>	<p>Use this to select a license under which you are sharing the data. The most common licensing options for open data sharing are listed. Please see the <a href="https://data.hdx.rwlab.org/about/license">DATA LICENSES</a> page at <a href="https://data.hdx.rwlab.org/about/license">https://data.hdx.rwlab.org/about/license</a> for a detailed explanation of each license. You may select the 'Multiple Licenses' if your dataset contains data that is licensed under more than one license. In this case, the various licenses should be made explicit in an additional metadata file if possible. You may also select 'Other' if you would like to type in your own license terms.</p> <p><b>COD</b></p> <ul style="list-style-type: none"> <li>• If the current license is defined by the <a href="#">COD FOD Registry Terms of Use</a> (Humanitarian use only) select "OTHER" from the choices (you can add the Humanitarian Use Only URL identified in the next metadata element).</li> <li>• Best practice is to try to gain agreement to have open data and to help data sources understand why open data is helpful to all (including themselves).</li> </ul>
<p><b>Define License</b></p>	<p>Use this field to state the terms and conditions under which you will only be available if you select 'Other' in the License field above.</p> <p>Example: Licensed under a <a href="#">Creative Commons AttributionNonCommercialShareAlike 4.0 International License</a> . This means that you can use this data as long as you attribute the source, you do not use the data for commercial purposes and you agree to share any derivatives works under the same license.</p> <p><b>COD</b> If you are unsure of the dataset license you should use chose "other" and leave it blank</p>
<p><b>Organisation</b></p>	<p>Use this field to select the organisation under which you are sharing the data. Your selection will be limited to the list of organisations that you are an administrator or editor of. The dataset will appear on the organisation page of the organisation that you select here (if it is public, see below).</p> <p><b>COD</b></p> <ul style="list-style-type: none"> <li>• All CODs should be registered as OCHA &lt;Country/Region&gt;.</li> <li>• FODs can be registered in different ways as the data sponsor or source may want to register it for ease of maintenance.</li> <li>• It is important that the dataset includes FOD in the title and tagged as an FOD so</li> </ul>

	operational partners are aware that this is the data set that is being used for operational purposes if there are multiple versions of a datasets theme.
<b>Visibility</b>	<p>Use this field to determine who can view your dataset. If you set the Visibility to 'Private', your dataset will not show up in search results. It will only be visible to users who are members of the organisation selected in Organisation above. Setting this field to "Public" will make the dataset accessible to anyone on the internet.</p> <p><b>COD</b></p> <ul style="list-style-type: none"> <li>Priority is to make CODs public, if data cannot be made public until an a disaster occurs datasets should be posted online and kept "private" so it only takes a toggle to make visible, if data must be requested create an HDX dataset and generic contact details or use the HDX contributor's button to receive reques.t</li> </ul>
<b>Tags</b> This metadata element is used in the search algorithm.	<p>Use the tags field to provide keywords for faceted searches of your dataset. Consider limiting the number of tags to 4 or 5. Use one word tags whenever possible. As you begin to type in a tag, you will be provided with a list of existing tags which you can choose from. It is better to select an existing tag from the list than to create your own. It is often worth the effort to also add the common synonyms for the tag. Add a 'geodata' tag if your dataset contains geographic data used in GIS systems. It is not necessary to add the same locations that you selected in the Locations field as tags.</p> <p>Example: cholera, health, wash, water, sanitation</p> <p><b>COD</b></p> <ul style="list-style-type: none"> <li>This is mandatory for CODs and FODs as Tags are used in the search algorithm. All CODs and FODs should include "COD" or "FOD" in the tag for easy discovery. It is also useful to include the COD themes- see list below for designation used in the COD FOD migration.</li> </ul>
Maintainer	<p>Use this field to give the name of the person who is responsible for the maintenance of the dataset. This information will not be displayed to users on the website, but is available on the HDX API. This selection determines who receives messages from other users who have questions or comments about the dataset.</p> <p>Example: Jane Doe</p>
Date of Dataset or Date Range	<p>Use the Date of Dataset field to specify the date that the data contained in the dataset was published. Alternatively, use the Date Range field to specify the date range that the data pertains to. You can only specify either the Date of Dataset or the the Date Range , not both.</p> <p>Example 1: 31/12/2014 (Date of Dataset) Example 2: 01/01/2014 31/12/2014 (Date Range)</p> <p><b>COD</b></p> <ul style="list-style-type: none"> <li>Consider this carefully, if the data has been cleaned or processed but is still the same data this should not be changed but the date of the new version should be noted in "Methodology" (e.g. the administrative boundaries or schools are still from 2013 but pcodes have been corrected or school names have been added but their locations have not been verified or updated - you would want clients to know it is still from this date and not reflective of more current data.</li> </ul>
Methodology	<p>Use this field to select the methodology that best describes the way the data in your dataset was collected. You can select 'Other' if you would like to specify a methodology that is not on the list of methodologies.</p> <p><b>COD</b></p> <ul style="list-style-type: none"> <li>The available choices are: census, sample data, direct observational data/anecdotal data, registry, other. for most CODs and FODs chose "other"</li> </ul>
Define Methodology	<p>Use this field to describe the methodology that was used to collect the data. This field will only be available if you select 'Other' in the Methodology field above.</p> <p>Example: This data has been collected as a combination of health ministry records and direct observation</p> <p><b>COD</b></p> <ul style="list-style-type: none"> <li>Add details of how the data was created, processing done to the data by OCHA or partners (e.g. pcodes or unique codes added, was updated through the IMWG or in partnership with a partner, if locations were collected by GPS or through satellite imagery, if population statistics have been estimated from historic census data, etc).</li> </ul>

	<ul style="list-style-type: none"> <li>Describe changes to the dataset so clients know if they should download the new version or not and the impact it will have on their systems. Include the date of the changes.</li> </ul> <p>Example:</p> <ul style="list-style-type: none"> <li>Based on satellite imagery taken on July 6, 2014.</li> <li>(2014-06-25) age and sex disaggregation added using national statistics</li> <li>Gathered through rapid assessments as well as monthly monitoring of market prices, climate, crop and livestock situations. For more information, visit <a href="http://www.faosomalia.org/fsnau">http://www.faosomalia.org/fsnau</a></li> <li>Education Facility Centres in Northern Uganda - based on GPS coordinates of the centers collected (in 2009, 2010, and 2011) by UNOCHA from different stake-holders working in Northern part of Uganda, such as cluster leads, humanitarian partners, and local government. The locations are re-verified by District Education Office (local government) before analysing and publishing Education Service Accessibility level against national standard.</li> </ul>
Caveats/ Comments	<p>Use this field to state any other pertinent information about the data. Pertinent information may include disclosures about any issues that you are aware of that affect the quality of the data you are sharing (e.g. the data from one of the locations in the dataset is an estimate from a previous survey because the location was inaccessible).</p> <p>Example: While the data is presented at admin level 3, the admin level 2 value was used in those instances where the data for admin level 3 was not available. This happened in 15% of admin level 3 locations, mainly in the south of the country</p> <p>COD</p> <ul style="list-style-type: none"> <li>Remember this is operational data and is “best available”.</li> <li>Data should be shared even if incomplete (e.g. does not have north part of country) so operational partners can start using it. Take advantage of this field to include notes about issues/challenges about the data so partners are aware, they also may be able to provide information/data/resources to improve the dataset or have another dataset the humanitarian community should examine.</li> </ul> <p>Example:</p> <ul style="list-style-type: none"> <li>(2014-09-16) Link to data updated. approximately 75 % completed by OSM, focus on urban area</li> <li>This list aggregates various data sources. The same locations might duplicated with different spelling. The data is in the process of being cleaned and will placed here at the end of June 2015</li> </ul>

### Resource metadata fields

Name	Use this field to provide a name for your dataset. The field will be populated with filename of the file by default
Note	Use this field to provide a concise description of the file (e.g. This file is a zip archive containing 10 excel
Format	<p>Use this field to specify the file format of the resource. If the resource is a zip file containing a single set of and it is generally ok not to change the default. worksheets) shapefiles, then you may specify 'ZIPPED SHAPEFILE' to enable the geopreview function on the resource.</p> <p>COD</p> <p>Be clear about the format of the data users want to know what software they will need or if it is in a format they have the skills to work with before they download it (especially when bandwidth is limited). If it is few excel files that are zipped say it is “excel” as any data can be zipped) exception to the rule:</p> <p>GEOPREVIEW FUNCTIONALITY - if you want your SHP to be spatially visualized on a map the format must be “ZIPPED SHAPEFILE” or “geojson”</p>

### Suggested list of COD FOD Tags

CFR_CATEGORIES	HDX_TAGS [English]	HDX_TAGS [French]
Admin Level 0 - 7 Boundaries	cod, administrative boundaries	cod, Limites administratives

Humanitarian Profile	cod, humanitarian profile, displacement	cod, Profile humanitaire
Canals	cod, hydrology	cod, Hydrologie
Water Courses	cod, hydrology	cod, Hydrologie
Water Bodies	cod, hydrology	cod, Hydrologie
Hydrology	cod, hydrology	cod, Hydrologie
Hypsography	cod, hypsography, elevation	cod, Courbes de niveau
Contour Lines	cod, hypsography, elevation	cod, Courbes de niveau
Spot Heights	cod, hypsography, elevation	cod, Courbes de niveau
Elevation Model	cod, hypsography, elevation	cod, Courbes de niveau
Settlements (villages, towns, cities)	cod, populated places	cod, Lieux peuplées
Populated Places	cod, populated places	cod, Lieux peuplées
Population Statistics	cod, population	cod, Statistiques de population
Aerodromes (airports, airfields, HLZs, etc.)	cod, transportation, infrastructure, airports	cod, Réseau du transport, infrastructure
Railways	cod, transportation, infrastructure	cod, Réseau du transport, infrastructure
Bridges	cod, transportation, infrastructure	cod, Réseau du transport, infrastructure
Ports	cod, transportation, infrastructure	cod, Réseau du transport, infrastructure
Roads	cod, transportation, infrastructure	cod, Réseau du transport, infrastructure
Railroads	cod, transportation, infrastructure	cod, Réseau du transport, infrastructure
Transportation Network	cod, transportation, infrastructure	cod, Réseau du transport, infrastructure
FOD		
Affected Persons Locations (Refugee/IDP Camps, etc.)	fod, affected persons locations, idp, refugee, camp, displacement	fod, Lieux des personnes affectées (camp PDI ou réfugiées, etc.)
Damage Assessments	fod, damage assessment	fod, Evaluation des dégâts
Education	fod, education,	fod, Education
Elections	fod, elections	fod, Elections
Food Security	fod, food security	fod, Sécurité Alimentaire
Health	fod, health	fod, Santé
Logistics	fod, logistics	fod, Logistique
Nutrition	fod, nutrition	fod, Nutrition
Protection	fod, protection	fod, Protection
Security	fod, security	fod, Sécurité
Water Sanitation Hygiene	fod, wash, watsan, sanitation, water, hygiene	fod, Eau hygiène et assainissement ,système sanitaire, eaux, hygiène
FOD Other *	fod, other	fod, autre
<p>*other - any other datasets that the humanitarian community agrees upon or uses in its processes (e.g. secondary data analysis for joint/harmonized assessments  For instance: banks, building footprints, parks, stadiums, land use, demographic data, etc</p>		