

3.1 Designing The Data Model

This section provides:

- An overview of data models and OpenStreetMap tagging, including presentation slides
 - Step-by-step instructions for tools to help build your data model
 - Step-by-step instructions for creating your data model
-

Overview

When beginning a mapping project, a data model will need to be created in order to determine what features are mapped and the details collected for each of those features. A **data model** defines what features are surveyed or mapped and what attributes are collected for each feature. If a project will upload data to OpenStreetMap, the data model should be designed to match OSM tagging.

Example data models:

- Uganda Refugee Crisis
- Ramani Huria

OSM doesn't work with layers or attribute tables, but tags. **Tags** are used in OSM to categorize features, and to add information that is useful for understanding of the map, planning, routing, and querying. Each tag consists of a key, and a value. Each map feature should have one or more tags such as:

- building=residential
- highway=primary
- amenity=school

In addition, each of these features can have an unlimited number of related attributes added into OSM as tags. For example, a building might have the following tags:

- building=commercial
- building:material=brick
- roof:material=metal
- shop=tailor

Creating a data model should be done in collaboration with all stakeholders in order to ensure that all necessary information is collected - it is a lot more difficult to revisit a location for mapping to collect additional information. At the same time, when designing a data model you should consider how much time it will take for a surveyor to complete data collection - every feature type, attribute, or question will add additional time needed.

Resources and Training Materials

This section features a selection of resources targeted at project managers, trainers, or even self-learners on the topic(s) outlined above.



Training Presentations

The following presentation(s) can be used to lead a training or workshop.

- Data models and tagging



Step-by-Step Guide

The following section is designed to serve as self-paced material that can be used both during trainings,

and by self-guided learners.

- Tools for creating your data model
- Creating your data model

Tools for creating your data model

HOT recommends using TagInfo and the OSM Wiki to search for existing OSM features when developing your data model. The following steps will provide an introduction to how to use these tools.

Skills and Technology Needed

- Computer with
- Internet connection
- OpenStreetMap Account
- Recommended: computer mouse

How to Use the OSM Wiki



The screenshot shows the OpenStreetMap Wiki page for 'Map Features'. On the left is a sidebar with a 'Wiki' logo and a list of links: Main Page, The map, Map Features, Contributors, Help, Blogs, Shop, Donations, Recent changes, Tools, What links here, Related changes, Special pages, Printable version, Permanent link, Page information, and Cite this page. The main content area has a header with 'English', 'Create account', and 'Log in' links. Below the header are tabs for 'Page', 'Discussion', 'Read', 'View source', and 'View history', followed by a search bar. The title 'Map Features' is prominently displayed. A section titled 'Available languages — Map Features' lists various languages including English, Spanish, Esperanto, and many others. Below this is a link to 'Other languages — Help us translate this wiki'. The main text explains that OpenStreetMap represents physical features on the ground using tags attached to its basic data structures (nodes, ways, and relations). It also mentions the free tagging system and the community's role in defining standards.

Figure 1: osm_wiki_map_features

1. Navigate to https://wiki.openstreetmap.org/wiki/Map_Features in a Google Chrome or Mozilla Firefox internet browser.
2. This page provides documentation on common, existing OSM features categorized by type. These tables contain keys and values, along with comments and sometimes pictures to help define the tag. Scroll through the tables to explore the tags described.
3. Search for a particular tag by using 'Ctrl+F' on your keyboard. As an example, search for the tag that should be used for hospitals. To do so, hit 'Ctrl+F' on your keyboard, type 'hospital' in the search bar, and hit enter. This will bring you to the appropriate tag for hospitals. *Note: there might be more than one appropriate tag for a search item. Scroll through the results until you find the appropriate tag.*

- ## How to use TagInfo

Figure 2: taginfo

- ## Creating your data model

Skills and Technology Needed

- 3

- Spreadsheet software, such as LibreCalc, Google Sheets, or Excel (recommended for structuring your data model)

Data model workflow

Designing a **data model** based on **OSM tagging** is typically defined by the following questions and workflow:

1. What is the focus of the data collection? Consider how the data will be used.
2. What features do you want to collect? Identify the focus of the data collection.
3. Where are you collecting data? Data models can differ based on their location.
4. What has been done before? Draft data model by leveraging similar models.
5. What tags exist for features?
 1. Verify tag status through OSM Wiki
 2. Verify tag usage through TagInfo
6. Use approved tags where possible in data model
7. Do all stakeholders agree on the data model? Review data model and incorporate feedback from project partners (adding, removing, or changing data model features can delay field data collection and decrease data quality!)

Create a list of all of the features that you want to collect. For example: buildings, water points, roads.

Feature
Buildings
Water Points
Roads

Go to OpenStreetMap wiki to search for the appropriate key for each feature, and value if there is only one value option.

Feature	Key	Value
Buildings	building	
Water Points	amenity	water_point
Roads	highway	

For features with multiple values, such as buildings, use the OSM Wiki page for that key as well as TagInfo to find appropriate values. These values should only be what is reasonable for your data collection. While it would be ideal to collect every building type in a city, your project might only be able to collect all school and hospital buildings. Note: values in your data model should make sense for the context of your geography. For example: hut is an appropriate value for buildings in Liberia, but not likely in Germany. Additionally, you may need to interpret an existing value type to best match the appropriate value for your region.

Feature	Key	Value
Buildings	building	residential, school, civic
Water Points	amenity	water_point
Roads	highway	primary, secondary, residential

Once you have the base tags for your features, you can decide on what attributes you want or can collect for each feature.

Feature	Key	Value
Buildings	building building:material building:levels	residential, school, civic

Feature	Key	Value
Water Points	roof:material	water_point
	amenity	
	status	
Roads	highway	primary, secondary, residential
	name	
	condition	
	surface	
	width	

Next, values can also be determined for each attribute key. These options can be determined using the OSM Wiki and TagInfo, or in some cases can be defined by the mapper - such as for numeric answers or names.

Feature	Key	Value
Buildings	building	residential, school, civic
	building:material	cement_block, brick, wood, mud
	building:levels	<i>numeric</i>
Water Points	roof:material	thatch, metal, concrete, plastic, tile
	amenity	water_point
	drinking_water	yes, no
Roads	highway	primary, secondary, residential
	name	<i>user defined</i>
	condition	excellent, good, poor
	surface	gravel, paved, dirt
	width	<i>numeric</i>

Once completed with your data model, this data model should be checked by your stakeholders for any gaps. Additionally, your project plan should allow flexibility so that this data model can be adjusted with field testing and consulting your mappers.

Collecting private data Private data should never be uploaded to OSM. However, some projects do require personal information to be collected. When this is the case, the data model can include unique non-OSM tags for private data that needs to be collected. When cleaning data following data collection, this private data can be kept in a full dataset before being removed. Once private data is removed, the dataset can be uploaded to OSM.