

Implementation Guidelines

TELANGANA OPEN DATA POLICY

Open Data Unit

DEPARTMENT OF INFORMATION TECHNOLOGY, ELECTRONICS & COMMUNICATIONS
GOVERNMENT OF TELANGANA

Revision History

Version	Description of Change	Author	Date
1.0	First Version	Open Data Unit	29-May-19
1.1	Second Version	Open Data Unit	15-Jan-2020

Table of Contents

INTRODUCTION	4
OPEN DATA POLICY: AN OVERVIEW	5
IMPLEMENTATION PROCESS.....	6
DATA PUBLISHING STANDARDS	7
METADATA STANDARDS	7
Title.....	7
Description	7
License.....	7
Tags	7
Sectors.....	7
Department.....	8
Author.....	8
Frequency.....	8
Temporal Coverage	8
Granularity	8
Contact Name.....	8
Contact Email	8
Public Access Level	8
Data Dictionary	8
FREQUENCY OF DATA UPDATING.....	9
IMPLEMENTING WEB SERVICES FROM DEPARTMENTS	9
DATA GLOSSARY	9
Revenue Department	9
CDMA, MAUD Department	10
LARGE DATASETS	10
CAPACITY BUILDING	10

Introduction

Data is a valuable resource and a strategic asset to the Telangana Government, its partners, and the public. Managing this data as an asset and making it available, discoverable, and usable—in a word, open—not only strengthens our democracy, and promotes efficiency and effectiveness in government, but also has the potential to create economic opportunity and improve citizens' quality of life.

The ultimate value of data can often not be predicted. That is why the Telangana Government has released a policy that instructs departments to manage their data—and information more generally—as an asset from the start and, wherever possible, release it to the public in a way that makes it open, discoverable, and usable.

Asset and value potentials of data are widely recognised at all levels. Data collected or developed through public investments, when made publicly available and maintained over time, their potential value could be more fully realised. There has been an increasing demand by the community that such data collected with the deployment of public funds should be made more readily available to all, for enabling rational debate, increase transparency better decision making, and use in meeting civil society and government needs.

Efficient sharing of data among data owners and inter- and intra-governmental agencies along with data standards and interoperable systems is the need of the hour. Hence, there was a need to formulate a policy on Data Sharing and Accessibility which could provide an enabling provision and platform for proactive and open access to the data generated through public funds available with various departments/organisations of the Government of Telangana.

The Telangana Government's Open Data Project—this collection of implementation guidelines, code, tools, and case studies—to help departments adopt the Open Data Policy and unlock the potential of government data. Project Open Data will evolve over time as a community resource to facilitate broader adoption of open data practices in government. Anyone—government employees, contractors, developers, the general public—can view and contribute.

Open Data Policy: An Overview

The Open Data Policy (ODP) is designed so as to apply to all shareable non-sensitive data available either in digital or analogue forms, but generated using public funds by various departments/subordinate offices/organisations/agencies of the Government of Telangana. The ODP is designed to promote data sharing and enable access to the Government of Telangana-owned data for state planning, development, and awareness.

The ODP aims to provide an enabling provision and platform for proactive and open access to the data generated by various Government of Telangana entities. The objective of this policy is to facilitate access to the Government of Telangana-owned shareable data (along with its usage information) in machine-readable form through a wide area network all over the country in a periodically updateable manner, within the framework of various related policies, acts and rules of the Government of India, thereby permitting a wider accessibility and usage by the public.

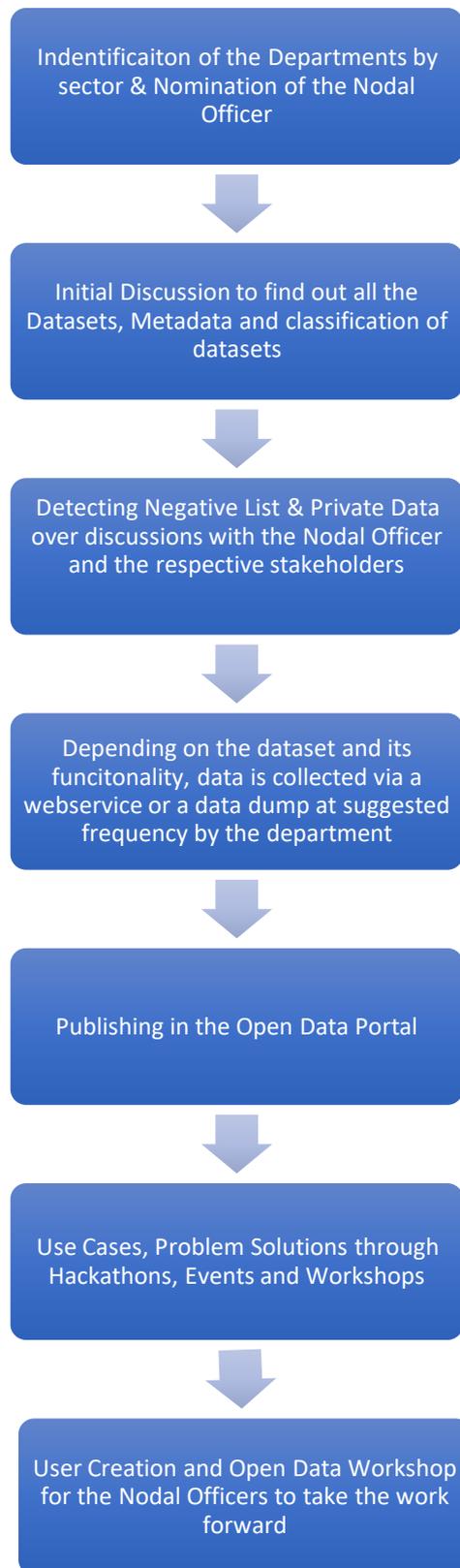
Different types of datasets generated both in geospatial and non-spatial form by different departments are supposed to be classified as shareable data and non-shareable data. Data management encompasses the systems and processes that ensure data integrity, data storage and security, including metadata, data security and access registers. The principles on which data sharing and accessibility need to be based include:

Openness, Flexibility, Transparency, Quality, Security and Machine-readability.

The ODP, once implemented, would lead to:

- Opening up of the information out of the government system
- Making available accurate, reliable and unbiased information
- Providing single data OGD platform for the state for data sharing
- Establishment of a platform to promote innovation in government applications
- Enhancing government transparency, accountability and public engagement
- Effective utilisation of government data by providing meaningful visual representations
- Enabling development of innovative applications around datasets or mash-up from multiple datasets, hence giving different perspectives to government data

Implementation Process



Data Publishing Standards

Post the initial discussion with the department about the data they hold with keen importance given to the data collection methods and context of the data collection, the Open Data Cell advises the department for the release of data in structured machine-readable formats such as a CSV or XLS.

Depending on the type of data provided by the departments, users can download the dataset in various formats such as csv, html, xls, json, xlsx, doc, docx, rdf, txt, jpg, png, gif, tiff, pdf, odf, ods, odt, tsv, tab, geojson, xml, zip. For specific cases, departments can provide web services, which can also be added directly as a dataset or a resource in the application.

Selection of the datasets is based on the department to share the information if it is in the negative list and has privacy concerns.

Timelines for data publishing shall be aligned with the department's decision to share the data.

Metadata

Metadata for the datasets may contain the one or more of the following entities

- Title
- Description
- License
- Tags
- Sectors
- Department
- Author
- Frequency
- Temporal Coverage
- Granularity
- Contact Name
- Contact Email
- Public Access Level
- Data Dictionary

There are other metadata that shall be included for datasets where such information is available.

Title

Title of the dataset should clearly provide insight into the what the dataset is about.

Description

Description of the data should explain the context of data collection method, data collected, the columns present in the data, and their units or specification to provide complete understanding of the dataset.

License

All the datasets on the ODP have the license of Open Government License, India. Except for a few cases where it is requested by the department for Non-Commercial License, e.g., weather data.

Tags

Tags are utilised to make it easy for users to search for the datasets or can be used to share the data with tags, e.g., mentioning the granularity 'Mandal Level Data' can provide users with all datasets with mandal-level data.

Sectors

Sectors describe the category to which the dataset belongs. One dataset can belong to multiple sectors.

Department

Source of the dataset, i.e., from which department the dataset is obtained. For datasets which are obtained from central government departments but pertains to Telangana State, the department should be specifically mentioned in the description of the dataset.

Author

Author is the section of the department which has collected the dataset, e.g., Department of Agriculture & Cooperation has multiple sections like TS Marketing, which collects the Daily Market Yard Prices.

Frequency

Stipulated time frame in which the dataset will be updated, as per discussion with the department.

Temporal Coverage

Temporal coverage is the time frame of the data, if the dataset pertains to the financial year or a quarter of the year, etc.

Granularity

Granularity is the level to which the dataset is present. Day wise, month wise, district level, mandal level, ward level, etc.

Contact Name

Name of the relevant person/team who can answer queries related to the dataset.

Contact Email

Email ID of the relevant person/team who can answer queries related to the dataset.

Public Access Level

Level of access of the dataset: Public, Restricted or Private.

Data Dictionary

Standards and information describing the contents, format, and structure of the dataset.

Data Standards the ODU use

Usage of globally accepted data standards makes the dataset more usable globally. ODU prefers the usage of such global data standards for its datasets if they are available. One such standard is Global Transit Feed Specification (GTFS), a standard developed by Google for Transit Datasets across the globe. ODU shall make the effort to update this document as and when other standards are used.

Global Transit Feed Specification (GTFS)

Utilizing the global standards for transit data has made it easier for the technology enthusiasts to develop solutions and emulate the same across the world for many scenarios through which the public will be benefitted enormously. Across the globe, transit information is being shared in the Open Domain by using Global Transit Feed Specification (GTFS) file formats developed by Google.

This specification defines the following files along with their associated content:

Filename	Required	Defines
agency.txt	Required	One or more transit agencies that provide the data in this feed
stops.txt	Required	Individual locations where vehicles pick up or drop off passengers
routes.txt	Required	Transit routes. A route is a group of trips that are displayed to riders as a single service

Filename	Required	Defines
trips.txt	Required	Trips for each route. A trip is a sequence of two or more stops that occur at a specific time
stop_times.txt	Required	Times that a vehicle arrives at and departs from individual stops for each trip
calendar.txt	Conditionally required	Dates for service IDs using a weekly schedule. Specify when service starts and ends, as well as days of the week where service is available. This file is required unless all dates of service are defined in calendar_dates.txt
calendar_dates.txt	Conditionally required	Exceptions for the service IDs defined in the calendar.txt file. If calendar.txt is omitted, then calendar_dates.txt is required and must contain all dates of service
fare_rules.txt	Optional	Rules for applying fare information for a transit organization's routes
shapes.txt	Optional	Rules for drawing lines on a map to represent a transit organization's routes
frequencies.txt	Optional	Headway (time between trips) for routes with variable frequency of service
transfers.txt	Optional	Rules for making connections at transfer points between routes
feed_info.txt	Optional	Additional information about the feed itself, including publisher, version and expiration information

Frequency of Data Updating

Frequency of the data updation for the dataset shall be adhered as requested by the department during the initial discussions with the individual departments.

Example:

1. Daily weather data is updated at a frequency of every 3 months, i.e., quarterly.
2. GTFS Dataset is updated at a frequency of every 3 months, i.e., quarterly.
3. Agricultural Area at the mandal level is updated half-yearly, based on cropping patterns.

Implementing Web Services from Departments

In cases where a department provides web services for data collection, the following parameters are taken into consideration:

1. Fields for all the data that is required, their formats
2. REST or SOAP API's and sample requests
3. Detailed description of all the methods in web service
4. If the web service can be made public

Open Data Unit creates a setup to collect data at regular intervals, store it in a database, and upload the data at a frequency based on the dataset.

Data Glossary

Data Glossary shall be made available for a few fields which are used across all the departments, such as district names, mandal names, etc., which can be standardised across all the departments. The rest of the fields, such as lake names or market yards, which are used by individual departments, shall be taken up on a case-by-case basis, and the data glossary for the individual cases shall be developed in due course.

Revenue Department

Data which contains the actual district names, mandal names, village names with IDs as decided by the Revenue Department after bifurcation of 10 districts to 33 districts. This shall be issued as a G.O. from the Revenue Department.

Sample data will be in the format below:

District Name	District Id	Revenue Division	Mandal Name	Mandal Id	Village Name	Village Id	Old District	Old Revenue Division
Bhadradi Kothagudem	###	Kothagudem	Kothagudem	###	Kothagudem	###	Khammam	Kothagudem
Bhadradi Kothagudem	###	Kothagudem	Kothagudem	###	Ramavaram	###	Khammam	Kothagudem

CDMA, MAUD Department

Data which contains the standardised names of urban local bodies, their grades, and revenue wards in each district with blocks, so that all the departments which work with urban data should be using this standardised dataset.

Large Datasets

In cases where the dataset is very large, the Open Data Unit shall divide the dataset into smaller chunks, so that users can download these in more easily. The division of the datasets can be done by time period, geography and data parameters.

Capacity Building

The Open Data Unit endeavours to conduct events with either departments and/or the user community and/or students to educate them on various facets of Open Data. For departments, the capacity building exercise would be about using the open data application, cleaning data, structures and formats of data in which it needs to be shared. They would also be trained on how to use the data for interesting visualisations using the tools that are freely available.