Implementation Guidelines

for
National Data Sharing and Accessibility Policy (NDSAP)

Ver. 2.1

April 2013

Data Centre and Web Services Division
National Informatics Centre
Department of Electronics and Information Technology
Ministry of Communications and Information Technology
Government of India

Revision History

Version	Description of Change	Author	Date
0.1	Initial Draft	National Informatics Centre, DeitY	10-Sept-2012
0.9	Incorporated suggestions received from the 1st Task Force meeting held on 17th September 2012	National Informatics Centre, DeitY	20-Sept-2012
1.0	Incorporated suggestions received from the 2nd Task Force meeting held on 1st October 2012	National Informatics Centre, DeitY	04-Oct-2012
1.1	Updated after the Data Controllers' Workshop held on 5th October 2012	National Informatics Centre, DeitY	08-Oct-2012
1.2	Incorporated feedbacks received from some Data Controllers	National Informatics Centre, DeitY	31-Oct-2012
2.0	Incorporated most of the feedbacks received from Data Controllers	National Informatics Centre, DeitY	07-Jan-2013
2.1	Updated the Auto-consumption of Datasets, Allowed Data Formats, Task Force Member List, NIC Team etc.	National Informatics Centre, DeitY	04-April-2013

Table of Contents

1.0   Introduction

Asset and value potentials of data are widely recognized 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 realized. There has been an increasing demand by the community, that such updated 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 National 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 ministries/departments/organizations of Government of India.

2.0   NDSAP: An Overview

The National Data Sharing and Accessibility Policy (NDSAP) is designed so as to apply to all sharable non-sensitive data available either in digital or analog forms but generated using public funds by various Ministries/Departments /Subordinate offices/organizations/ agencies of Government of India. The NDSAP policy is designed to promote data sharing and enable access to Government of India owned data for national planning, development and awareness.

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

Different types of data sets generated both in geospatial and non-spatial form by different ministries/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-readable.*

NDSAP once implemented would lead to:

1. Opening up of the information out of the Government System
2. Makin available of the Accurate, Reliable and Unbiased information
3. Providing single data portal for the country for data sharing
4. Establishment of a platform to promote innovation in government applications
5. Enhancing government Transparency, Accountability and Public Engagement
6. Effective utilization of Government data by providing meaningful visual representations
7. Enabling development of Innovative Applications around datasets or mash-up from multiple datasets hence giving different perspectives to government data

Complete text of Policy can be accessed at Annexure I, while Gazette Notification for NDSAP can be accessed at http://www.india.gov.in/allimpfrms/alldocs/16473.pdf.

3.0   Open Government Data

"A dataset is said to be open if anyone is free to use, reuse, and redistribute it – Open Data shall be machine readable and it should also be easily accessible."

Government collects processes and generates a large amount of data in its day-to-day functioning. But a large quantum of government data remains inaccessible to citizens, civil society, although most of such data may be non-sensitive in nature and could be used by public for social, economic and developmental purposes.

These data need to be made available in an open format to facilitate use, reuse and redistribute; it should be free from any license or any other mechanism of control. Opening up of government data in open formats would enhance transparency and accountability while encouraging public engagement. The government data in open formats has a huge potential for innovation building various types of Apps, mash-ups and services around the published datasets.

Notification of NDSAP mandates government departments to proactively open up data. NDSAP in India is applicable to all entities of Government Setup.

3.1 Identification of Datasets

As per policy each department has to prepare it's Negative List. The datasets which are confidential in nature and are in the interest of the country's security in not opening to the public would fall into the negative list. This list would need to be compiled and sent to the Department of Science and Technology within six months.

All other datasets which do not fall under this negative list would be in the Open List. These datasets would need to be prioritized into high value datasets and non-high value datasets.

As per the NDSAP, within a year all the datasets need to be published on the Data Portal, data.gov.in and within the first three months at least 5 high value datasets need to be published.

Government data generated through following processes and events:

1. Primary Data e.g. Population Census, Education Census, Economic Survey, etc.
2. Processed/Value Added Data e.g. Budget, Planning, etc.
3. Data Generated through delivery of Government Services e.g. Income Tax Collection, MNREGA wage distribution etc.

The data which are contributed to the data portal have to be in the specified open data format only. The data have to be internally processed to ensure that the quality standard is met i.e. accuracy, free from any sort of legal issues, privacy of an individual is maintained and does not compromise with the National security.

While prioritizing the release of datasets, one should try to publish as many high value datasets. Though each department shall have its own criterion of high value and low value datasets, generally High value data is governed by following Principles (for details on this please refer to Annexure – III):

1. Completeness
2. Primary
3. Timeliness
4. Ease of Physical and Electronic Access
5. Machine readability
6. Non-discrimination
7. Use of Commonly Owned Standards
8. Licensing
9. Permanence
10. Usage Costs

3.2 Data Formats

NDSAP recommends that data has to be published in open format. It should be machine readable. Though there are many formats suitable to different category of data. Based on current analysis of data formats prevalent in Government it is proposed that data should be published in any of the following formats:

- CSV (Comma separated Values)
- XLS (spread sheet- Excel)
- ODS/ODT (Open Document Formats for Spreadsheet/Text)
- XML (Extensive Markup Language)
- RDF (Resources Description Framework)
- KML (Keyhole Markup Language used for Maps)
- GML (Geography Markup Language)
- RSS/ATOM (Fast changing data e.g. hourly/daily)

Tim Berners-Lee had classified data into single star to five star categories based on formats for data. Please refer to Annexure II.

4.0   Data Portal India

Data Portal India has been set up at http://data.gov.in to provide collated access to datasets published by different government departments in open format. It also provides a search & discovery mechanism for instant access to desired datasets. Data Portal also has a rich mechanism for citizen engagement. Besides enabling citizens to express their need for specific datasets or apps, it also allows them to rate the quality of datasets, seek clarification or information from respective data controller.

Data Portal has a backend data management system which can be used by government departments to publish their datasets through a predefined workflow. They shall also have a dashboard to see the current status on their datasets, usage analytics as well as feedback and queries from citizens at one point.

Data Portal also has a Communities component which facilitates forming of communities around datasets, domain of interest such as agriculture, education, health, or it could be app developers' community or even of data journalists. This shall give first hand input to development community for building new components, apps. It shall also give input to departments as what kind of datasets is more useful and accordingly prioritize the release of the datasets. The key features are detailed as below:

- **Open Source Driven** –Developed completely using Open Source Stack, facilitating cost saving in terms of software and licenses and also provisioning community participation in terms of further development of product with modules of data visualization, consumption, APIs to access datasets etc.

- **Metadata** – Datasets shall be published along with standard metadata along with controlled vocabularies on government sectors, jurisdictions, dataset types, access mode etc. Besides facilitating easy access to datasets, this shall be extremely useful in the future for federation/integration of data catalogs.

- **Social Media Connect**–To support wider reach and dissemination of datasets, anyone can share the information about any dataset published on the portal with his/her social media pages on a press of a click.

- **Citizen Engagement** –Portal has a strong component of Citizen Engagement. Citizens can express their views as well as rate the datasets w.r.t three aspects (Quality, Accessibility and Usability) on the scale of 5. They can also embed the datasets in their blogs or web sites. Facility to contact the data controllers is also available on the Portal.

- **Community Collaboration**– Citizens with specific interest can build communities and discuss online. Portal facilitates the communities to open up online forums, blogs and discussions around various datasets, apps available on the portal. It also provides a platform to express and discuss the kind of datasets & apps they would like to have. It shall also give input to departments as what kind of datasets is more useful and accordingly prioritize the release of the datasets.

5.0 NDSAP Implementation in Ministries/Departments

In order to implement NDSAP, the Ministries/Departments of Government of India have to undertake the following activities:

a. Nominate Data Controller
b. Setup NDSAP Cell
c. Identify Datasets
d. Publish Datasets on Data Portal India
e. Prepare Negative List
f. Create Action Plan for regular release of datasets on the Data Portal India
g. Monitor and Manage the Open Data Programme of the Department

5.1 Data Controller

As per the directive from the Cabinet Secretary, a senior officer at the level of Joint Secretary or above is to be nominated as the Data Controller or Nodal Officer for the Department/Organization. The responsibility of Data Controller is as follows:

1. Lead the open data initiative of department.
2. Take initiative to release as many datasets as possible on proactive basis.
3. Identify the High Value Datasets and schedule their release on Data Portal
4. Prepare the Negative List for the Department as per the directions in NDSAP.
5. Ensures that the datasets being published are in compliance with NDSAP through a workflow process.
6. Periodically monitor the release of datasets as per predefined schedule
7. Take relevant action on the feedback/suggestion received from the citizen for the datasets belonging to the Ministry/Department.
8. Ensure the correctness of his contact details on the Data Portal by sending a mail to ndsap@gov.in, in case of any change.
9. Take action on Suggestions on new datasets made by public on data Portal.

5.2 NDSAP Cell

In order to implement NDSAP each Department would establish a NDSAP Cell. The size of the cell would vary from Department to Department and would depend on the quantum of datasets to be published. The NDSAP Cell would be responsible for:

- Prepare Negative List of datasets and communicate to DST within Six Months
- Prepare a schedule of datasets to be released in next one year
- Extend Technical Support for Preparation of datasets, conversion of formats etc.
- Monitor and manage the Open data initiative in their respective Ministry/ Department and ensure quality and correctness of the data
- Work out an open data strategy to promote proactive dissemination of datasets
- Institutionalize the creation of datasets as part of routine functioning

NDSAP Cell shall be headed by Data Controller, who could be assisted by number of Data Contributors. NDSAP Cell shall have professionals from data analyst, visualization and programming domain. The policy mentions that budgetary provisions and appropriate support for data management for each department/organization would be necessary.

5.3 Data Contributor

In order to cater to the contribution of the datasets from offices/organization under the Ministries/Departments, the Data Controller can nominate a number of Data Contributors who would be responsible in contributing the datasets along with their metadata. Using the web based DMS, each data contributor would be able to contribute the data as per the given metadata format (elaborated at section 7) which is based on the Dublin Core Standards. The contributed datasets would be approved by the Data Controller as the case may be.

Data Contributor could be an officer of the Ministry/Department who would be responsible for his/her unit/division. The responsibilities of the Data Contributor are as follows:

- Responsible for ensuring quality and correctness datasets of his/her unit/division
- Preparing and contributing the metadata in the predefined format for the datasets

6.0 Publishing & Management of Datasets

Contribution of datasets is through a simple web based Dataset Management System https://datacms.nic.in. Datasets contributed are processed through a predefined workflow, ensuring compliance with government policies. Data Controllers nominated by government ministries or department are authorized to publish datasets in open format on Data Portal.

Data Controllers of the Ministry/Department has the facility to create any number of Contributors for contributing Datasets/Apps for their Ministry/Department. Once the Contributor is created by the Data Controller, a mail is sent to the mail id of the contributor. The Contributor then can login and contribute datasets along with its metadata for further approval by the Controller. However, the responsibility on the relevancy and quality of datasets published on the Data Portal rests with Data Controller.

6.1 Registration for Creating Datasets/Apps

Data Controllers:

Nominated Data Controllers having NIC mail Ids can directly register with the Portal by "Signing In" using his/her NIC Email credentials, i.e. same login id and password through a web based Dataset Management System https://datacms.nic.in.

However, Data Controllers who do not have NIC mail Id, will get their login credentials through mail at their email id provided in the nomination letter.

The Data Controllers may need to validate their Ministry/Department through the interface if shown after they login.

Data Contributors:

Datasets/apps are contributed by the Data Contributors by going to the web-based interface available at https://datacms.nic.in and using their credentials which are sent to their mail id.

6.2 Contribute, Moderate/Approve & Publish

- Datasets along with the metadata once contributed by the contributor goes through the predefined workflow to the Data Controller who in turn ensures that it is in compliance with the NDSAP Policy. After the approval of the Data Controller the dataset is ready to be published on the Data Portal.
- The NDSAP PMU pushes the datasets from staging area to the production area.
- Any dataset which is published can be modified/ edited but it would pass through the workflow before publishing the edited dataset and archiving the earlier dataset.
- Any published dataset can only be deleted by the Data Controller who is the owner of the dataset of the department.

6.3 Consumption of Datasets from eGov Applications

Metadata discovery is a way to fetch datasets and metadata information from different eGov applications. Desired eGov applications have to define the source and type of tool to be used to fetch dataset and metadata from their servers. The data.gov.in tool would crawl at source at specified time to fetch metadata/datasets.

Data Source Formats for Metadata Discovery

Programming languages or Scripts or SQL queries can be used to write some programs to fetch the data from the relational database or some other sources to generate the desired metadata and dataset information. The data.gov.in tool can fetch the data catalog information with the help of the programming language. The language passes pre-defined metadata information to the data import tool to create the data catalogs automatically without any human interface. Datasets and related metadata may be created from a variety of different sources including:

- Relational databases
- CSV files or Spreadsheets
- XML files
- Web Services or data in JSON format

Relational Databases

All most all the eGov applications use a relational database to store their reports and MIS data. Stored procedures or Database Triggers can be written to periodically export the data in to a dataset in CSV or XML format which can be consumed by data.gov.in data import tool to create the dataset catalog.

CSV Files or Spreadsheets

CSV files or Spreadsheets are the most efficient and common source of data.gov.in data import tool. eGov applications can specify the location from which data import tool has to fetch the dataset and metadata. Metadata and dataset can be put in a zip file with predefined naming convention. With the help of the tool data.gov.in can traverse designated location and extract the metadata and dataset in CSV files or Spreadsheets to convert it into data catalogs.

XML files

Structured XML files are one of the good sources for the data import. eGov applications can pass the metadata information and the details about the metadata in a predefined XML file. As defined by the eGov application data.gov.in data import tool can traverse the designated location of the XML at predefined time to fetch the data and convert it into searchable human readable data catalogs.

Web Services or Data in JSON Format

Web service can be used as a stream of data or services. data.gov.in supports both REST and SOAP based web service over HTTP. Using REST based web services provides more flexibility to share data over XML as well as JSON format. eGov application has to submit the definition of the fields to be exposed by the web services which need to be mapped with the metadata information to create the catalog. At the given period of time with the help of metadata mapping the tool will import the data from the source apps and will convert to data catalogs.

6.4 View & Respond to Queries on Published Datasets

Citizens can browse, search and access the datasets on the Portal. Citizens also have the option to send their queries and feedbacks about the published datasets.

This feedback would be available on the dash board of the Data Controller to take further necessary action.

6.5 Respond to Suggestions for new Datasets

The Data Portal has a strong Citizen Engagement feature built in. While browsing through the catalogue of datasets, if one is not able to find the dataset which is of interest to him then he can request for the same through suggestions module. Suggestions already made for particular datasets are displayed and one can also endorse the same. The suggested list i.e. the requirement for new datasets is sent to the respective department's data controller. This would facilitate the data controller to prioritize his release of datasets on the portal. They are expected to send in their response on the same.

6.6 Review Analytics & Plan

The Dash Board of the Data Controller metrics would be available for the datasets contributed by all the contributors of that department. Feedback related to datasets would also be available along with the suggestions from citizen with respect to the requirement of new datasets. This feature would facilitate him to watch the analytics and accordingly plan his course of action.

6.7 DOs for Data Contribution and Approval

- Prioritize the release of datasets, categorize the type of access granted and publish as many high value datasets.
- Contribute datasets which are in Open List and do not fall under Negative List.
- Ensure that the datasets being published are in compliance with NDSAP through a workflow process. Data should be collected at the source, and includes the original information collected by the government, details on how the data was collected and the original source documents recording the collection of the data.
- Prepare and contribute the metadata in predefined format for the datasets. The key metadata elements are Title, Description, Sector/Sub-Sector, Dataset Jurisdiction, Keywords, Access Method, Reference URLs, Data Group Name, Frequency and Policy Compliance. All the metadata elements must be filled with

utmost quality and ease of use.

- Pricing of data, if any, would be decided by the data owners and as per the government policies.

- Ensure that data being contributed to the data portal are in machine readable or in specified open data format only. The advisable formats as described in Section 3.2 above.

- Ensure that the data is as complete as possible, reflecting the entirety of what is recorded about a particular subject and is de-normalized before uploading.

- Priority should be given to data whose utility is time sensitive. Real time information updates would maximize the utility the public can obtain from this information.

- Replace any Not Available, Not Reported or missing values in the data with 'NA'.

- Metadata that defines and explains the raw data should be included as well, along with formulas and explanations for how derived data was calculated.

- Keywords must be defined in data catalog to make it search friendly.

- Provide link to the reference documents (if any) or website for detailed information and explanation on the method of calculation or source of data.

- Read the process manual below for step by step procedure for upload of data.

Process Manual

- Prioritize the release of datasets and take relevant action on the basis of feedbacks and suggestions received on the Data Portal from citizen's pertaining to the Ministry/ Department.

- Ensure the correctness of login details on the Data Portal by sending a mail to ndsap@gov.in, in case of any change.

6.8 DON'Ts for Data Contribution and Approval

- Contribute datasets which are confidential in nature and are in the interest of the country's security.

- Data has not been externally processed to ensure that the quality standards are met i.e. accuracy, free from any sort of legal issues, privacy of an individual is

maintained and does not compromise with the National security.

- The imposition of 'Terms of Service', attribution requirements, restrictions on dissemination and so on acts as barriers to public use of data. Use available data without restrictions on use as part of the public domain.

- Cost imposed on the public for access. Imposing fees for access skews the pool of who is willing (or able) to access information.

- Hand written notes on paper are very difficult for machines to process. Scanning text via Optical Character Recognition (OCR) results in many matching and formatting errors. Information shared in the widely used PDF format is very difficult for machines to parse. Hence, data in these formats are not allowed.

- Data in non-Unicode formats should be avoided.

- Data or Dataset File Name should not contain special characters e.g. @, %, $, &, etc. or missing values.

- Any explanation, including the method of calculation or source of data need not be provided in data file to be attached in the web form.

7.0   Metadata Elements and their Description

1. **Title** (Required)**:**

   A unique name of the Dataset viz. Current Population Survey <Year>, Consumer Price Index <Year> etc. This field will be used to populate the data catalog; the catalog will be sorted on this field. It may be noted that if the title is not unique within the entire catalog, you will be asked to change it by the system during contribution.

2. **Description** (Required)**:**

   Please provide a detailed description of the Dataset e.g., an abstract determining the nature and purpose of the catalog. Facility is there to provide text of around 6000 characters.

3. **Sector/Sub-Sector** (Required)**:**

   Choose the sectors(s)/sub-sector(s) those most closely apply(ies) to your dataset.

4. **Dataset Jurisdiction** (Required)**:**

   This is a required field to identify the exact location or area to which the dataset caters to viz. entire country, state/province, district, city etc.

5. **Keywords** (Required)**:**

   It is a list of terms, separated by commas, describing and indicating at the content of the dataset. Example: rainfall, weather, monthly statistics,

6. **Access Method (Required):**

   This could be "Upload' or "Link".
   - **Upload**: It would ask to choose the dataset from the local directory to upload to the central repository.
   - **Link**: It would ask to provide the exact URL of the dataset to access that in a single click. It would also ask to enter file size and file format.URL to the high level repository/resource, in which this dataset or tool resides at the agency. It can also be called as the one point link location to that particular dataset.

7. **Reference URLs:**

   URLs for the technical or usage documentation for this Dataset. This may include description to the study design, instrumentation, implementation, limitations, and appropriate use of the dataset or tool. In the case of multiple documents or URLs, please delimit with commas or enter in separate lines.

8. **Dataset Group Name:**

   This field allows agencies to provide a Group Name to closely related datasets in order to show that they may be presented as a group or a set.

9. **Frequency:**

   It mentions the time interval over which the data inside the dataset is collected/ updated on a regular basis (one-time, annual, hourly, etc.). To be selected from the dropdown menu.

10. **Policy Compliance**

    This field is to indicate if this dataset is in conformity with the National Data Sharing and Access Policy of the Govt. of India.

8.0 NDSAP Project Management Unit (PMU)

The main activities of NDSAP, PMU would be to manage the Data Portal, provide Technical Advice to the departments, handhold for dataset contribution as well as capacity building of Data Contributors and Data Controllers.

8.1 Management of Data Portal

Data Portal India would be managed and hosted at the National Data Centre, NIC Headquarters adhering to the Guidelines of the Government and Data security policies. The architecture would be scalable and of high availability.

8.2 Technical Advice

Departments would be provided with technical advice with respect to preparation of datasets, contribution of datasets, explanation of metadata and the entire workflow of data publishing, feedback management etc.

8.3 Capacity Building

Two types of training modules both as offsite and onsite models have been envisaged. Each module would be for the duration of 2-3 days. The logistics and venue for the onsite training would be the responsibility of the host organization. The modules would be:

- Awareness and Sensitization Module – for Data Controller & other senior officers of the Ministries/Departments
- Data Contribution Module – hands-on training for contributing datasets to the data portal, provide advisory on conversion of data to digital format to Data Contributors and Members of NDSAP Cell.

8.4 Contact Details

For further details regarding implementation of NDSAP in the Ministry/Department you may contact:

> NDSAP PMU (Programme Management Unit)
> National Informatics Centre, DeitY
> Room No. 375, 3rd Floor
> A Block, CGO Complex, Lodi Road New Delhi – 110003
> Phone: 011-24363692/011-24305296, Fax: 011-24368854
> Email: ndsap@gov.in

Annexure – I

National Data Sharing and Accessibility Policy (NDSAP) – 2012

1.   Preamble

1.1   Asset and value potentials of data are widely recognized 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 realized. 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, better decision making and use in meeting civil society needs. Principle 10 of the United Nations Declaration on Environment and Development (Rio de janeiro, June 1992), stated

"………..each individual shall have appropriate access to information concerning the environment that is held by public authorities ……. and the opportunity to participate in the decision-making process. States shall facilitate and encourage public awareness and participation by making information widely available."

Section 4(2) of the Right to Information Act, 2005 reads

"It shall be a constant endeavor of every public authority to take steps in accordance with the requirements of clause (b) of sub-section (1) to provide as much information suomotu to the public at regular intervals through various means of communication, including internet, so that the public have minimum resort to the use of this Act to obtain information"

1.2   The principles on which data sharing and accessibility need to be based include: Openness, Flexibility, Transparency, Legal Conformity, Protection of Intellectual Property, Formal Responsibility, Professionalism, Standards, *Interoperability, Quality, Security, Efficiency, Accountability, Sustainability and Privacy.*

1.3   A large quantum of data generated using public funds by various organizations and institutions in the country remains inaccessible to civil society, although most

of such data may be non-sensitive in nature and could be used by public for scientific, economic and developmental purposes. The National Data Sharing and Accessibility Policy (NDSAP) is designed so as to apply to all sharable non-sensitive data available either in digital or analog forms but generated using public funds by various Ministries / Departments /Subordinate offices / organizations / agencies of Government of India. The NDSAP policy is designed to promote data sharing and enable access to Government of India owned data for national planning and development.

2. Definitions

2.1 Data –Data means a representation of information, numerical compilations and observations, documents, facts, maps, images, charts, tables and figures, concepts in digital and/or analog form.

2.2 Data Archive – A place where machine-readable data are acquired, manipulated, documented, and distributed to others for further analysis and consumption.

2.3 Data Generation– Initial generation/collection of data or subsequent addition of data to the same specification.

2.4 Data set – A named collection of logically related features including processed data or information.

2.5 Geospatial Data – All data which is geographically referenced

2.6 Information – Processed data

2.7 Metadata – The information that describes the data source and the time, place, and conditions under which the data were created. Metadata informs the user of who, when, what, where, why, and how data were generated. Metadata allows the data to be traced to a known origin and know quality.

2.8 Negative List – Non-sharable data as declared by the departments/ organizations

2.9 **Restricted Data** –Data which are accessible only through a prescribed process of registration and authorization by respective departments / organizations.

2.10 **Sensitive Data**– Sensitive data as defined in various Acts and rules of the Government of India.

2.11 **Sharable Data** – Those data not covered under the scope of negative list and non-sensitive in nature

2.12 **Standards –** Any application that embeds data handling functions (e.g., data collection, management, transfer, integration, publication, etc.) and operates on data in a manner that complies with data format and data syntax specifications produced and maintained by open, standards bodies.

3. **Need for the Policy**

Evidence-based Planning of socio-economic development processes rely on quality data. There is a general need to facilitate sharing and utilization of the large amount of data generated and residing among the entities of the Government of India. This would call for a policy to leverage these data assets which are disparate. The current regime of data management does not enable open sharing of Government owned data with other arms of the government nor does it expect proactive disclosure of sharable data available with data owners. Such regimes could lead to duplication of efforts and loss of efficiency of planning of activities focused on national development. Efficient sharing of data among data owners and inter and intra governmental agencies and with public calls for data standards and interoperable systems. Hence, National Data Sharing and Access Policy aims to provide an enabling provision and platform for providing proactive and open access to the data generated through public funds available with various   departments / organizations of Government of India.

4. **Objectives**

The objective of this policy is to facilitate the access to Government of India owned shareable data and information in both human readable and machine readable forms through a network all over the country in a proactive and periodically updatable manner, within the framework of various related policies,

Acts and rules of Government of India, thereby permitting a wider accessibility and use of public data and information.

5.   Scope of this Policy

The National Data Sharing and Accessibility Policy will apply to all data and information created, generated, collected and archived using public funds provided by Government of India directly or through authorized agencies by various Ministries / Departments /Organizations / Agencies and Autonomous bodies.

6.   Benefits of the data sharing policy

6.1   Maximizing Use – Ready access to government owned data will enable more extensive use of a valuable public resource for the benefit of the community.

6.2   Avoiding Duplication – By sharing data the need for separate bodies to collect the same data will be avoided resulting in significant cost savings in data collection.

6.3   Maximized integration – By adopting common standards for the collection and transfer of data, integration of individual data sets may be feasible.

6.4   Ownership information – The identification of owners for the principal data sets provide information to users to identify those responsible for implementation of prioritized data collection programs and development of data standards.

6.5   Better decision-making – Data and information   facilitates making important decisions without incurring repetitive costs. Ready access to existing valuable data is essential for many decision making tasks such as protecting the environment, development planning, managing assets, improving living conditions, national security and controlling disasters.

6.6   Equity of access – Amore open data transfer policy ensures better access to all bonafide users.

7.   Data Classification

Different types of data sets generated both in geospatial and non-spatial form by different ministries /departments are to be classified as shareable data and non-shareable data. The types of data produced by a statistical system consist of derived statistics like national accounts statistics, indicators like price index, data bases from census and surveys. The geospatial data however, consists primarily of satellite data, maps, etc. In such a system, it becomes important to maintain standards in respect of metadata, data layout and data access policy. All departments / ministries will prepare the negative list within one year of the notification of the policy, which will be periodically reviewed by the oversight committee.

8.   Types of Access

8.1   Open Access – Access to data generated from public funding should be easy, timely, user-friendly and web-based without any process of registration / authorization.

8.2   Registered Access – Data sets which are accessible only through a prescribed process of registration / authorization by respective departments / organizations will be available to the recognized institutions / organizations / public users, through defined procedures.

8.3   Restricted Access – Data declared as restricted, by Government of India policies, will be accessible only through and under authorization.

9.   Technology for Sharing and Access

A state-of-the-art data warehouse and data archive with online analytical processing (OLAP) capabilities, which includes providing, a multi-dimensional and subject oriented view of the database needs to be created. This integrated repository of data portals of various ministries / departments as a part of data.gov.in, will hold data and this repository over a period of time will also encompass data generated by various State Governments and UTs. The main features of the data warehouse need to include:

a) User friendly interface

b) Dynamic / pull down menus

c) Search based Report

d) Secured web access

e) Bulletin board

f) Complete Metadata

g) Parametric and Dynamic report in exportable format

10. Legal framework

Data will remain the property of the agency/department/ ministry/ entity which collected them and reside in their IT enabled facility for sharing and providing access. Access to data under this policy will not be in violation of any Acts and rules of the Government of India in force. Legal framework of this policy will be aligned with various Acts and rules covering the data.

11. Pricing

Pricing of data, if any, would be decided by the data owners and as per the government policies. All Ministries / Departments will upload the pricing policy of the data under registered and restricted access within three months of the notification of the policy. A broad set of parameters would be standardized and provided as guidelines for the use of data owners.

12. Implementation

a) The Department of Science & Technology serving the nodal functions of coordination and monitoring of policy through close collaboration with all Central Ministries and the Department of Information Technology by creating data.gov.in through National Informatics Centre (NIC).

b) All sharable data will be made available on 'as-is where-is' basis.

c) Detailed implementation guidelines including the technology and standards for data and metadata would be brought out by Department of Information Technology, Government of India.

d) All the data users who are accessing/using the data shall acknowledge the ministry/department in all forms of publications.

e) All Ministries/Departments will upload at least 5 high value data sets on data.gov.in within three months of the notification of the policy.

f) Uploading of all remaining data sets should be completed within one year

g) Thereafter, all data sets are to be uploaded regularly every quarter.

h) data.gov.in will have the metadata and data itself and will be accessed from the portals of the departments/ministries.

i) The metadata in standardized formats is to be ported on data.gov.in which enables data discovery and access through departmental portals. All metadata will follow standards and will minimally contain adequate information on proper citation, access, contact information, and discovery. Complete information including methods, structure, semantics, and quality control/assurance is expected for most datasets.

j) Government will design and position a suitable budgetary incentive system for data owners for increasing open access to the sharable data.

k) An oversight committee will be constituted for facilitating the implementation of the policy and its provisions thereof

l) Department of Information Technology will constitute a coordination committee for implementation.
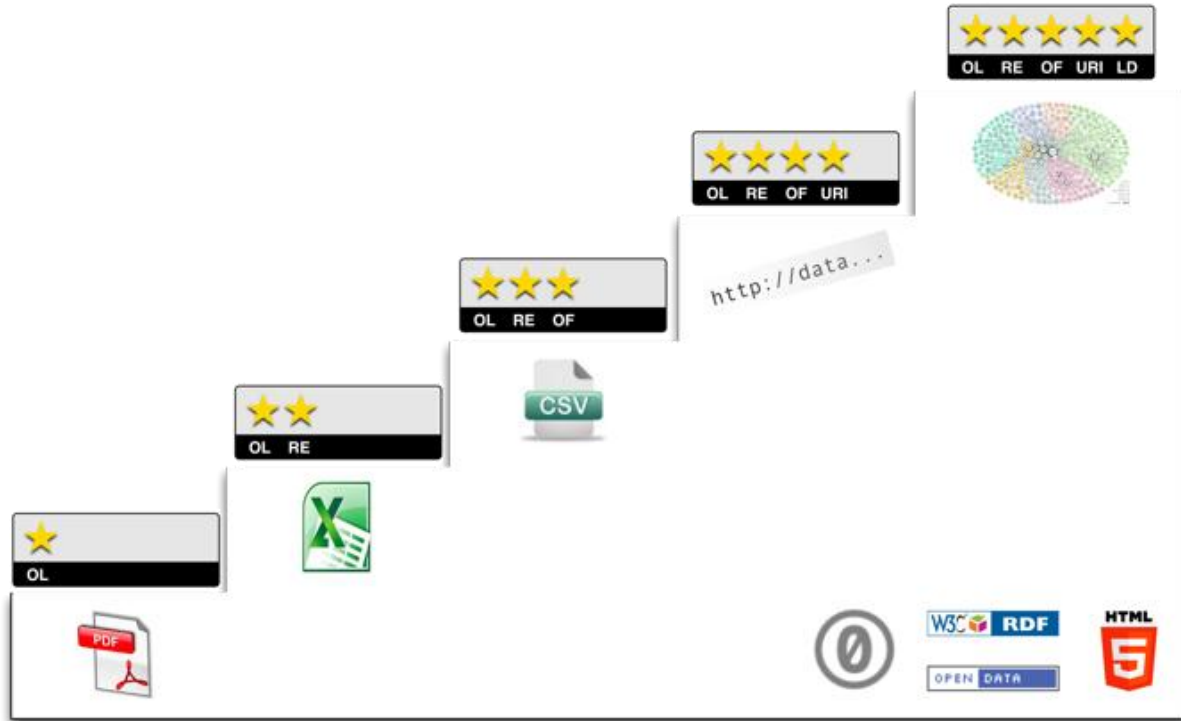
13. Budget Provisions

The implementation of National Data Sharing and Access Policy is expected to entail expenditures for both data owners and data managers for analog to digital conversion, data refinement, data storage, quality up-gradation etc. Budgetary provisions and appropriate support for data management for each department / organization by Government of India would be necessary.

14. Conclusion

14.1 While policies provide official mandate, facilitation of optimum accessibility and usability of data by the implementers pre-suppose a trajectory of proper organization of data, with access services and analysis tools that provide the researchers and stakeholders with added value. For data to be reused, it needs to be adequately described and linked to services that disseminate the data to other researchers and stakeholders. The current methods of storing data are as diverse as the disciplines that generate it. It is necessary to develop institutional repositories, data centers on domain and national levels that all methods of storing and sharing have to exist within the specific infrastructure to enable all users to access and use it.

14.2 National Data Sharing and Access Policy aims at the promotion of a technology-based culture of data management as well as data sharing and access. It opens up, proactively, information on available data, which could be shared with civil society for developmental purposes, their price details if any, and methods for gaining access to registered and restricted use. The policy has limited its scope to data owned by the agencies, departments/ Ministries and entities under the Government of India and forms a statement of the Government of India of its commitment to transparency and efficiency in governance. Department of Science & Technology will continue the process of evolving the policy further, keeping in tune with technological advancements and the National requirements and enrolling the State Governments.

Five Star Open Data
(Source http://5stardata.info/)

Tim Berners-Lee, the inventor of the Web and Linked Data initiator, suggested a 5 star deployment scheme for Open Data. Here, we give examples for each step of the stars and explain costs and benefits that come along with it.

Where the abbreviations stands for:

> OL – Open License
> RE – Readable (Human & Machine)
> OF – Open Format
> URI – Universal Resource Identifier
> LD – Linked Data

The cost and benefit of these open data levels are detailed below:

★ Available on the web (whatever format) but with an open licence, to be called as Open Data

As a consumer.....
- You can look at it.

- You can print it.
- You can store it locally (on your hard drive or on an USB stick).
- You can enter the data into any other system.
- You can change the data as you wish.
- You can share the data with anyone you like.

As a publisher.....
- It's simple to publish.
- You do not have explain repeatedly to others that they can use your data.

Note: It's great to have the data accessible on the Web under an open license (such as PDDL, ODC-by or CC0), however, the data is locked-up in a document. Other than writing a custom scraper, it's hard to get the data out of the document

★★ Available as machine-readable structured data (e.g. excel instead of image scan of a table)

As a consumer, you can do all what you can do with ★ Web data & additionally…..
- You can directly process it with proprietary software to aggregate it, perform calculations, visualise it, etc.
- You can export it into another (structured) format.

As a publisher ...
- It's still simple to publish

Note: Splendid! The data is accessible on the Web in a structured way (that is, machine-readable), however, the data is still locked-up in a document. To get the data out of the document you depend on proprietary software.

★★★as (2) plus non-proprietary format (e.g. CSV instead of excel)

As a consumer, you can do all what you can do with ★★ Web data and additionally…..
- You can manipulate the data in any way you like, without being confined by the capabilities of any particular software.

As a publisher .....
- You might need converters or plug-ins to export the data from the proprietary format.
- It's still rather simple to publish.

Note: Excellent ! The data is not only available via the Web but now everyone can use the data easily. On the other hand, it's still data on the Web and not data in the Web.

★★★★ All the above plus, Use open standards from W3C (RDF and SPARQL) to identify things, so that people can point at your stuff using an URI

As a consumer, you can do all what you can do with ★★★ Web data and additionally…...
- You can link to it from any other place (on the Web or locally).
- You can bookmark it.
- You can reuse parts of the data.
- You may be able to reuse existing tools and libraries, even if they only understand parts of the pattern the publisher used.
- Understanding the structure of an RDF "Graph" of data can be more effort than tabular (Excel/CSV) or tree (XML/JSON) data.
- You can combine the data safely with other data. URIs are a global scheme so if two things have the same URI then it's intentional, and if so that's well on it's way to being 5 star data!

As a publisher ...
- You have fine-granular control over the data items and can optimise their access (load balancing, caching, etc.)
- Other data publishers can now link into your data, promoting it to 5 star!
- You typically invest some time slicing and dicing your data.
- You'll need to assign URIs to data items and think about how to represent the data.
- You need to either find existing patterns to reuse or create your own.

Note: Wonderful! Now it's data in the Web. The (most important) data items have a URI and can be shared on the Web. A native way to represent the data is using RDF. However, other formats such as Atom can be converted/mapped, if required.

★★★★★ All the above, plus: Link your data to other people's data to provide context

As a consumer, you can do all what you can do with ★★★★ Web data and additionally…...
- You can discover more (related) data while consuming the data.
- You can directly learn about the data schema.
- You now have to deal with broken data links, just like 404 errors in web pages.

- Presenting data from an arbitrary link as fact is as risky as letting people include content from any website in your pages. Caution, trust and common sense are all still necessary.

As a publisher .....
- You make your data discoverable.
- You increase the value of your data.
- You own organisation will gain the same benefits from the links as the consumers.
- You'll need to invest resources to link your data to other data on the Web.
- You may need to repair broken or incorrect links.

Note: Brilliant! Now it's data, **in** the Web **linked to** other data. Both the consumer and the publisher benefit from the network effect

Ten Principles for Opening Up Government Information
(Source: http://sunlightfoundation.com/policy/documents/ten-open-data-principles/)

In October 2007, 30 open government advocates met in Sebastopol, California to discuss how government could open up electronically stored government data for public use. Up until that point, the federal and state governments had made some data available to the public, usually inconsistently and incompletely, which had whetted the advocates' appetites for more and better data. The conference, led by Carl Malamud and Tim O'Reilly and funded by a grant from the Sunlight Foundation, resulted in eight principles that, if implemented, would empower the public's use of government held data.

Sunlight Foundation has updated and expanded upon the Sebastopol list and identified **ten principles** that provide a lens to evaluate the extent to which government data is open and accessible to the public. The list is not exhaustive, and each principle exists along a continuum of openness. The principles are ***completeness, primacy, timeliness, ease of physical and electronic access, machine readability, nondiscrimination, use of commonly owned standards, licensing, permanence and usage costs***.

1. Completeness
Datasets released by the government should be as complete as possible, reflecting the entirety of what is recorded about a particular subject. All raw information from a dataset should be released to the public, except to the extent necessary to comply with federal law regarding the release of personally identifiable information. Metadata that defines and explains the raw data should be included as well, along with formulas and explanations for how derived data was calculated. Doing so will permit users to understand the scope of information available and examine each data item at the greatest possible level of detail.

2. Primacy
Datasets released by the government should be primary source data. This includes the original information collected by the government, details on how the data was collected and the original source documents recording the collection of the data. Public dissemination will allow users to verify that information was collected properly and recorded accurately.

3. Timeliness

Datasets released by the government should be available to the public in a timely fashion. Whenever feasible, information collected by the government should be released as quickly as it is gathered and collected. Priority should be given to data whose utility is time sensitive. Real time information updates would maximize the utility the public can obtain from this information.

4. Ease of Physical and Electronic Access

Datasets released by the government should be as accessible as possible, with accessibility defined as the ease with which information can be obtained, whether through physical or electronic means .Barriers to physical access include requirements to visit a particular office in person or requirements to comply with particular procedures (such as completing forms or submitting FOIA requests). Barriers to automated electronic access include making data accessible only via submitted forms or systems that require browser oriented technologies (e.g., Flash, Javascript, cookies or Java applets). By contrast, providing an interface for users to download all of the information stored in a database at once (known as "bulk" access) and the means to make specific calls for data through an Application Programming Interface (API) make data much more readily accessible. (An aspect of this is "findability," which is the ability to easily locate and download content.)

5. Machine readability

Machines can handle certain kinds of inputs much better than others. For example, hand written notes on paper are very difficult for machines to process. Scanning text via Optical Character Recognition (OCR) results in many matching and formatting errors. Information shared in the widely used PDF format, for example, is very difficult for machines to parse. Thus, information should be stored in widely used file formats that easily lend themselves to machine processing. (When other factors necessitate the use of difficult to parse formats, data should also be available in machine friendly formats.) These files should be accompanied by documentation related to the format and how to use it in relation to the data.

6. Non-discrimination

"Non-discrimination" refers to who can access data and how they must do so. Barriers to use of data can include registration or membership requirements. Another barrier is the use of "walled garden," which is when only some applications are allowed access to data. At its broadest, non-

discriminatory access to data means that any person can access the data at any time without having to identify him/herself or provide any justification for doing so.

7. Use of Commonly Owned Standards

Commonly owned (or "open") standards refers to who owns the format in which data is stored. For example, if only one company manufactures the program that can read a file where data is stored, access to that information is dependent upon use of the company's processing program. Sometimes that program is unavailable to the public at any cost, or is available, but for a fee. For example, Microsoft Excel is a fairly commonly used spreadsheet program which costs money to use. Freely available alternative formats often exist by which stored data can be accessed without the need for a software license. Removing this cost makes the data available to a wider pool of potential users.

8. Licensing

The imposition of "Terms of Service," attribution requirements, restrictions on dissemination and so on acts as barriers to public use of data. Maximal openness includes clearly labeling public information as a work of the government and available without restrictions on use as part of the public domain.

9. Permanence

The capability of finding information over time is referred to as permanence. Information released by the government online should be sticky: It should be available online in archives in perpetuity. Often times, information is updated, changed or removed without any indication that an alteration has been made. Or, it is made available as a stream of data, but not archived anywhere. For best use by the public, information made available online should remain online, with appropriate version tracking and archiving over time.

10. Usage Costs

One of the greatest barriers to access to ostensibly publicly available information is the cost imposed on the public for access—even when the cost is de minimus. Governments use a number of bases for charging the public for access to their own documents: the costs of creating the information; a cost recovery basis (cost to produce the information divided by the expected number of purchasers); the cost to retrieve information; a per-page or per-inquiry cost; processing cost; the cost of duplication etc.

Most government information is collected for governmental purposes, and the existence of user fees has little to no effect on whether the government gathers the data in the first place. Imposing fees for access skews the pool of who is willing (or able) to access information. It also may preclude transformative uses of the data that in turn generates business growth and tax revenues.

Annexure – IV

Frequently Asked Questions (FAQ)

Q. **What is the aim and objective of National Data Sharing and Accessibility Policy (NDSAP)?**

A. NDSAP aims to provide an enabling provision and platform for proactive and open access to the data generated by various Government of India entities.

The objective of this policy is to facilitate access to Government of India owned shareable data (along with its usage information) in machine readable form through a wide area network all over the country in a periodically updatable manner, within the framework of various related policies, acts and rules of Government of India, thereby permitting a wider accessibility and usage by public.

Q. **What is the scope of NDSAP?**

A. The National Data Sharing and Accessibility Policy will apply to all data and information created, generated, collected and archived using public funds provided by Government of India directly or through authorized agencies by various Ministries / Departments / Organizations / Agencies and Autonomous bodies.

Q. **What are the principles of data sharing and accessibility?**

A. The principles on which data sharing and accessibility need to be based include: Openness, Flexibility, Transparency, Quality, Security and Machine-readable.

Q. **Which Ministry / Department is performing the nodal function of coordination and monitoring of this Policy?**

A. The Department of Science & Technology serving the nodal functions of coordination and monitoring of policy through close collaboration with all Central Ministries and the Department of Electronics and Information Technology by creating data.gov.in through National Informatics Centre (NIC).

Q. **How datasets are identified?**

A. As per NDSAP, every Department has to identify datasets by the following categories:

Negative List: The datasets, which are confidential in nature and would compromise to the county's security if made public, are put into this list.

Open List: This list comprises of datasets which don't fall under negative list. These datasets shall be prioritized into high value datasets and non-high values datasets.

Q. What are the formats in which datasets should be published?

A. NDSAP recommends that datasets has to be published in an open format. It should be machine readable. Considering the current analysis of data formats prevalent in Government, it is proposed that data should be published in any of the following formats:

- CSV (Comma separated values)
- XLS (Spread sheet - Excel)
- ODS/~~OTS~~ (Open Document Formats for Spreadsheets)
- XML (Extensive Markup Language)
- RDF (Resources Description Framework)
- KML (Keyhole Markup Language used for Maps)
- GML (Geography Markup Language)
- RSS/ATOM (Fast changing data e.g. hourly/daily)

Q. What are the classifications of data?

A. Different types of datasets generated both in geospatial and non-spatial form by Ministries / Departments shall be classified as shareable data and non-shareable data.

The derived statistics like national accounts statistics, indicators like price index, databases from census and surveys are the types of data produced by a statistical mechanism. However, the geospatial data consists primarily of satellite data, maps, etc. In such a mechanism, it becomes important to maintain standards in respect of metadata, data layout and data access policy.

Q. What are the budget provisions in this Policy?

A. The implementation of National Data Sharing and Access Policy is expected to entail expenditures for both data owners and data managers for analog to digital conversion, data refinement, data storage, quality up-gradation etc. Budgetary provisions and appropriate support for data management for each department / organization by Government of India would be met from their internal budget for the formation of NDSAP cell.

Q. What activities are to be undertaken up by the Ministries/Departments to implement NDSAP?

A. In order to implement NDSAP, the Ministries/Departments of Government of India have to undertake the following activities:

- Nominate Data Controller, at the level of Joint Secretary or above
- Setup NDSAP Cell
- Identify Datasets
- Prepare Negative List
- Publish Datasets on Data Portal India
- Create Action Plan for regular release of datasets on the Data Portal India
- Monitor and Manage the Open Data Program of the Department

Q. Who shall be nominated as Data Controller?

A. As per the directive from the Cabinet Secretary, a senior officer at the level of Joint Secretary or above is to be nominated as Data Controller or Nodal Officer for the Department / Organization.

Q. What is the main role of a Data Controller & what are his/her responsibilities?

A. The main **role & responsibilities of a Data Controller are as follows:**

- Lead the open data initiative of Department / Organization.
- Take initiative to release as many datasets as possible on proactive basis.
- Identify the High Value Datasets and schedule their release on Data Portal.
- Prepare the Negative List for the Department as per the directions in NDSAP.
- Ensures that the datasets being published are in compliance with NDSAP through a workflow process.
- Periodically monitor the release of datasets as per predefined schedule.
- Take relevant action on the feedback/suggestion received from the citizen for the datasets belonging to the Ministry / Department.
- Ensure the correctness of his contact details on the Data Portal by sending a mail to ndsap@gov.in, in case of any change.
- Take action on Suggestions on new datasets made by public on Data Portal.

Q. What should be the size of NDSAP Cell?

A. In order to implement NDSAP, each Department would establish a NDSAP Cell. The size of the cell would vary from Department to Department and would depend on the quantum of datasets to be published.

Q. What are the responsibilities of NDSAP Cell?

A. The NDSAP Cell chaired by Data Controller would be responsible for the following activities:

- Prepare Negative List of datasets and communicate to DST within Six Months.

- Prepare a schedule of datasets to be released in next one year.
- Extend Technical Support for Preparation of datasets, conversion of formats etc.
- Monitor and manage the open data initiative in their respective Ministry/ Department and ensure quality and correctness of the data.
- Work out an open data strategy to promote proactive dissemination of datasets.
- Institutionalize the creation of datasets as part of routine functioning.

Q. What shall be the organizational structure of NDSAP Cell?

A. NDSAP Cell shall be headed by Data Controller, who could be assisted by number of Data Contributors. NDSAP Cell shall have professionals from data analyst, visualization and programming domain.

Q. What is the role of Data Contributor?

A. In order to cater to the contribution of the datasets from offices/organization under the Ministries / Departments, the Data Controller can nominate a number of Data Contributors who would be responsible in contributing the datasets along with their metadata.

Data Contributor could be an officer of the Ministry / Department who would be responsible for his/her unit/division.

The responsibilities of the Data Contributor are as follows:

- Responsible for ensuring quality and correctness datasets of his/her unit/division.
- Preparing and contributing the metadata in the predefined format for the datasets.

Q. How data is contributed on Data Portal India (http://data.gov.in/)?

A. Datasets are contributed through a web-based Data Management System **https://datacms.nic.in**, which has a simple interface. Data Controllers can directly register with the Data Portal India through Data Management System using their NIC EMAIL credentials.

Data Controllers, without having NIC EMAIL account, can obtain the login credentials (user name and password) on their existing non NIC EMAIL ID mentioned in the nomination letter.

Datasets along with the metadata once contributed by the contributor goes through the predefined workflow to the Data Controller who in turn ensures that it is in compliance with the NDSAP Policy.

The NDSAP PMU at NIC headquarters pushes the datasets from staging area to the production area.

Q. What is the role of NDSAP Project Management Unit (PMU)?

A. The main activities of NDSAP, PMU would be to manage the Data Portal, provide Technical Advice to the departments, handhold for dataset contribution as well as capacity building of Data Contributors and Data Controllers.

Q. What is metadata?

A. Data about the data is called metadata. Information about the datasets being published thru a standard structure comprising of controlled vocabularies on government sectors, dataset types, jurisdictions, access mode, etc. Apart from facilitating easy access to data, it is useful for federation & integration of data catalogs.

Q. What are metadata Elements and their usage?

A. **Title (Required)**: A unique name of the Dataset viz. Current Population Survey <Year>, Consumer Price Index <Year> etc.

Description (Required): Provide a detailed description of the Dataset e.g., an abstract determining the nature and purpose of the catalog.

Sector & Sub-Sector (Required): Choose the sectors(s)/sub-sector(s) those most closely apply (ies) to your dataset.

Dataset Jurisdiction (Required): This is a required field to identify the exact location or area to which the dataset caters to viz. entire country, state/province, district, city etc.

Keywords (Required): It is a list of terms, separated by commas, describing and indicating at the content of the dataset. Example: rainfall, weather, monthly statistics.

Access Method (Required): This could be "Upload a Dataset" or "Single Click Link to Dataset".

Reference URLs: This may include description to the study design, instrumentation, implementation, limitations, and appropriate use of the dataset or tool. In the case of multiple documents or URLs, please delimit with commas or enter in separate lines.

Dataset Group Name: This field allows agencies to provide a Group Name to closely related datasets in order to show that they may be presented as a group or a set.

Frequency: It mentions the time interval over which the data inside the dataset is collected/ updated on a regular basis (one-time, annual, hourly, etc.).

Policy Compliance: This field is to indicate if this dataset is in conformity with the National Data Sharing and Access Policy of the Govt. of India.

Q. What are the best practices for dataset contribution?

A. Some of the best practices for dataset contribution are as follows:

- Data should be stored in widely used file formats that are suitable for machine processing.
- Released dataset should clearly reflect "what is recorded about a particular subject".
- Timely release of datasets is one of the important factors to maximize the utility of information people can obtain.
- Data should be provided in freely available formats which can be accessed without the need for a software license.
- Data should be de-normalized.

Annexure – V

Constitution of the Task Force

Refer the Office Memorandum no. 21(1)/2013-Pers dated 2nd April 2013; following is the updated Task Force consisting of the following officers for formulation of Implementation Guidelines for National Data Sharing and Accessibility Policy (NDSAP):

1. **Dr. Y. K. Sharma**   **……. Chairman**
   Director General, National Informatics Centre

2. **Dr. (Smt) Shefali Sushil Dash**
   Deputy Director General, National Informatics Centre

3. **Shri Sanjay Singh Gahlout**
   Deputy Director General, National Informatics Centre

4. **Dr. R. Siva Kumar**
   Head NRDMS & NSDI, Department of Science & Technology

5. **Dr. C. Muralikrishna Kumar**
   Senior Advisor (CIT&I), Planning Commission

6. **Shri Rajiv Rai**
   Jt. Secretary, & CWO, Department of Personnel & Training

7. **Shri P. C. Mohanan**
   Dy. Director General, Ministry of Statistics and Programme Implementation

8. **Smt. Sudha Midha**
   Addl. Director General, Ministry of Water Resources

9. **Shri Pankaj Srivastava**
   Director, Ministry of Corporate Affairs

10. **Shri Gaurav Dwivedi**
    Director (eGovernance), Department of Electronics & Information Technology

11. **Representative**
    Department of Administrative Reforms and Public Grievances (DARPG)

12. **Ms. Neeta Verma**
    Senior Technical Director, National Informatics Centre

13. **Ms. Alka Mishra**   **……. Convener**
    Senior Technical Director, National Informatics Centre

Annexure – VI

Data Portal India Team

- **Ms. Neeta Verma, HoG and Head of Data Portal India**

- **Ms. Alka Mishra, Sr. Technical Director**

- **Mr. V. S. Mohan Das, Technical Director**

- **Mr. Durga Prasad Misra, Principal Systems Analyst**

- **Mr. Narender Jain, Systems Analyst**

- **Mr. Varun Gupta, Scientific/Technical Assistant**