Sikkim Open Data Acquisition and Accessibility Policy (SODAAP) – 2014

1. Preamble

1.1 In his budget speech delivered on April 25,2013, the Hon'ble Chief Minister of Sikkim, expressing his vision of a state-of-the-art data management system, made the following statement:

"Sikkim is at the threshold of entering into a new category of states that are seen as good models to be emulated due to high growth rate and good governance. However, on many occasions, our arguments and success stories lost most of their sheen due to lack of corresponding data. In order to be able to spur further growth in all areas of human development, the case for holistic acquisition of data, research and dissemination is being felt."

1.2 Open Data presents opportunities for public service transformations by giving users more power to self-serve. It creates a platform for more informed public debate. This in turn means the public is better equipped to hold their government accountable and also to be more informed about public services in their area. 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 states that

"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 suo motu 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."

1.3 The policy will be a driver for economic growth in the building of a knowledge society as envisaged by the Chief Minister. A new market for public service information will thrive if data is freely available in a standardized format for use and re-use. Open Data across government and public services would allow a market in comparative analytics, information presentation and service improvement to flourish. This new market will attract talented entrepreneurs and skilled employees, creating high value-added services for citizens, communities, non-profit organizations and public service providers, delivering auxiliary jobs and driving demand for skills.

1.4 The Sikkim Open Data Acquisition and Accessibility Policy (SODAAP) is designed so as to apply to all sharable non-sensitive data available either in digital(machine readable) or analog(human readable) forms but generated using public funds by various Ministries/Departments/Subordinate offices/organizations/agencies of Government of Sikkim. The SODAAP policy is designed both to provide authentic data to buttress the achievements of the Government of Sikkim and to gather data on key metrics to be able to spur growth in all the areas of human development.

2. Definitions

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

2.2 **Data Center** – 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 the 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 sources and the time, place, and conditions under which the data was created. Metadata informs the user of who, when, what, why, and how data were generated. Metadata allows the data to be traced to a known origin and a known quality.

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

2.9 **Open 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.

2.10 **Sensitive Data** – Sensitive Data refers to personal data or information of a person which consists of information relating to:

   a. Password;

b. Financial information such as bank account or credit card or debit card or other payment instrument details;

c. Physical, physiological and mental health condition;

d. Sexual orientation;

e. Medical records and history;

f. Biometric information;

g. Any detail relating to the above clauses as provided to third party for providing service;

h. Any of the information received under above clauses from third party for processing, stored or processed under lawful contract or otherwise

2.11 **Sharable Data** – Those data not covered under the scope of negative list and is 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 used.

3. Objectives

The following are the objectives of the policy as directed by the Hon'ble Chief Minister:

a. Create a **SODAA (Sikkim Open Data Acquisition and Accessibility) web portal and to host in Sikkim State Data Centre**, Gangtok where authentic data from all Government bodies will reside;

b. Design a mechanism to update the Data Bank regularly;

c. Enable formulation of reports and evaluation studies to empower public policy decisions; and

d. Provide access to all Departments and Institutions, both public, private and **citizens worldwide**, to this web portal.

4. Scope of the policy

The Sikkim Open Data Acquisition and Accessibility Policy will apply to all data and information created, generated, collected and archived using public funds provided by Government of Sikkim directly or through authorized agencies by various ministries/departments/organizations/agencies and autonomous bodies

5. Guidelines for classification of data

5.1 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 in the negative list.

5.2 All other datasets which do not fall under the negative and sensitive list would be in the Open List and includes the following types of data sets:

a. Primary Data e.g. Number of Beds in Hospitals, Doctor-to-nurses ratio, Number of Teachers in primary school etc.

b. Processed/Value Added Data e.g. Budget, Planning etc.

c. Data generated through delivery of Government Services e.g. Literacy Rate, Average Life Expectancy etc.

d. Geospatial data consisting of primarily satellite data, maps etc.

6. Technology for sharing and access

A state-of-the-art SODAA web portal will be developed and it will be hosted in Sikkim State Data Centre with online analytical processing (OLAP) capabilities needs to be created. This data centre is an integrated repository of data portals of various ministries / departments will hold data and this repository over a period of time. Some of the main features of the data bank (web portal) may include the following:

    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

7. Legal framework

Data will remain the property of the agency/department/ ministry/ entity which collected them and the same will be used 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 Sikkim in force (such as IPR, Copy Right, IT Rules, etc.). Legal framework of this policy will be aligned with various acts and rules covering the data.

8. Implementation

9.1 The Department of Information Technology will serve as the nodal authority to coordinate and monitor of policy through close collaboration with all State Government Departments and agencies. Department of Information Technology will allocate sufficient server space in Sikkim State Data Centre (SDC) to facilitate operationalizing this policy through an appropriate web-based portal.

9.2 All Departments will assess and identify incumbents for Data Controller, SODAA Cell, Data Contributors and other associated members within their Department for coordinating with Department of Information Technology on a regular basis.

9.3 All Departments will identify and prepare a comprehensive list of data sets maintained by them. Within 5(five) months from the date of notification of this SODAA policy, they will create a structure for each dataset with metadata in a standardized format. All metadata will follow standards within 3 months from the date of issue of SODAA Policy Notification and will minimally contain adequate information on proper citation, access and contact information (see Annexure I).

9.4 All Departments will prepare the negative list within six months of the notification of the policy, which will be periodically reviewed by the oversight committee. (See 9.8)

9.5 All sharable data will be made available, by the Departments, to web-based solution on an 'as-is where-is' basis.

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

9.7 The metadata in standardized formats will be hosted by the web-based solution, which enables data discovery and access through Departmental portals. The web-based solution will have the metadata and data itself and will be accessed from the portals of the Departments.

9.8 An oversight committee will be constituted for facilitating the implementation of the policy, monitoring the progress for the implementation and ensure it is completed in a time-bound fashion.

9.9 An oversight committee will periodically track the performance of the various Departments. In this direction, an annual self-assessment report will be generated by each Department followed by the assessment of the oversight committee.

9.10 Department of Information Technology will constitute a SODAA Project Management Unit (PMU) for supporting the implementation. The PMU will work closely with all the Departments for the implementation of the policy.

9.11 All the data users who are accessing/using the data shall acknowledge the Department in all forms of publications.

9.12 The state e-Governance Apex Committee may function as the oversight committee.

9. Roles and Responsibilities

10.1 **Data Controller** - 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:

a. Lead the open data initiative of Department

b. Take initiative to release as many datasets as possible on proactive basis

c. Identify the Open Datasets and schedule their release on Data web portal

d. Prepare the Negative List for the Department as per the directions in SODAAP

e. Ensures that the datasets being published are in compliance with SODAAP

f. Periodically monitor the release of datasets as per predefined schedule

g. Take relevant action on the feedback/suggestion received from the citizen for the datasets belonging to the Ministry/Department within one week from the receipt of the feedback/suggestion.

h. Ensure the correctness of his contact details on the Data Center by sending a mail to *<email id>*, in case of any change.

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

a. Prepare negative list of datasets

b. Prepare a schedule of datasets to be released in next one year

c. Extend technical support for preparation of datasets, conversion of formats etc.

d. Monitor and manage the open data initiative in their respective Ministry/ Department and ensure quality and correctness of the data

e. Work out an open data strategy to promote proactive dissemination of datasets

f. Institutionalize the creation of datasets as part of routine functioning

SODAAP Cell shall be headed by Data Controller, who could be assisted by number of Data Contributors. SODAAP Cell shall have professionals from data analyst, visualization and programming domain including State Government IT personnel's.

10.3 **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. Each data contributor would be contributing the data as per the given metadata format (elaborated in Annexure II). 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:

a. Responsible for ensuring quality and correctness datasets of his/her unit/division

b. Preparing and contributing the metadata in the predefined format for the datasets

10.4 **SODAAP Project Management Unit (PMU) -** The main activities of SODAAP PMU would be to manage the Data web portal, provide technical advice to the Departments, handhold for dataset contribution as well as capacity building of Data Contributors and Data Controllers.

a. Management of Data Center – Web Portal would be managed and hosted at *Sikkim State Data Centre, gangtok* adhering to the guidelines of the Sikkim Government and data security policies. The architecture would be scalable and of high availability

b. 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.

c. Capability Building – The following would be the two training modules:

1. Awareness and Sensitization Module – For Data Controller & other senior officers of the Ministries/Departments

2. 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 SODAAP cell

10. Annexures

Annexure – I

Meta Data elements and their description

<To be put down after finalization of the technicalities>