

**Background Paper**  
**for formulation of**  
**National Data Sharing and Accessibility Policy (NDSAP)**

**1. CLASSIFICATION OF DATA**

**Definition Paradigm:** Data and Information have gained high significance for developmental planning in knowledge societies. Civil societies seek open access to such data and information generated with public funds for planning developmental processes. On the other hand, defence sensitivity requirements demand the restriction of access and availability of sensitive data. With growing levels of terrorism and the powerful use of technologies by non-state actors, providing free access to data and information is a challenge faced by Nation States.

National Data Sharing and Accessibility Policy of India envisages a migration from one paradigm to another. The current classification of data of sharing is based on “Open Series Data” model. In this process, any data not specifically included in the “Open Series Dataset” remains inaccessible for public use.

The Government of India has accorded approval to the changed paradigm of migrating towards “Negative List” of data rather than definition of an open Data Series Model

Such Negative Lists could be based on features rather than nature and type.

Classification of Data based on feature and negative list for exclusion is the new paradigm.

Data owners and sources may therefore need to define and classify their data based on “features” and “exclusion principle” for preparing a negative list within a defined time frame.

For Geo-spatial data, the current mechanism of NSDI (jointly managed by DST and DoS) and for other types, DST will serve as a Nodal Point.

**Technical Issues:** Primary data generated by various agencies and sources become useful for the Civil Society needs only after value additions and generation of meta data. Continuous updating of data is also a necessity. For example, National Remote Sensing Imagery Data at a resolution of 1M may provide an accuracy level of 2 M. On the other hand, map products of SOI provide an accuracy level of 12.5 Meter at 1:50,000 resolution. Data generated by SOI enjoy the merit of Ground truthing and high reliability. In order to harmonize resolutions of map products with the accuracy and resolution of imagery policy, SOI may need to gather map data at a resolution of 1:8000. There is a technical issue in updating such data infrastructure. Further, in order to implement the National Data Sharing and Accessibility Policy standardization and interoperability issues as well as digitization of legacy data are essential.

Such infrastructure requirements call for strategic alliances and partnerships between primary data producers and other service providers. They also catalyse technology partnerships among various stake holders.

**Institutional and Managerial Issues:** Institutions generating the data and data managers are required to balance the need to restrict the use of data on security considerations with civil society and scientific community requirements for the same data, generated at the cost of public funds. Non-sensitive data of limited spatial resolution and time legacies should be, made quickly available for wider use and permitted access to registered users.

**Financial Issues:** Data generation and continuous updating do require financial support in the form of budgetary allocations or priced transactions. The policy framework should permit suitable financial support mechanisms either through budgetary support or rational pricing provisions.

**Legal Issues:** National and International agreements directly affect the data-access and data-sharing practices. Copyright and IPR considerations may lead to certain restrictions. Security considerations preclude access and sharing of some kinds of data. However, denial of data access leads to denial of possible benefits arising from the use of the data. Therefore, data access could be provided to users based on stated purposes and the public good considerations. Registered users for certain types of data for legitimate and legally permitted use could be provided data access through a policy framework. Since the purpose of use becomes an important consideration in the data sharing and access policy, legal protections against the non- legitimate use would become necessary.

**Operational Issues:** Nature of data, tools, technology infrastructure, structural and functional control over data and data use, standardization of data would raise a large number of operational issues while formulating data sharing and access policy, which have to be addressed.

## **2. GUIDELINES FOR CLASSIFICATION OF DATA**

General principles for providing access and enabled sharing arise from the potentials for public and social good associated with such type of data which could be shared without compromise to national security, IPR and other confidentiality considerations. Classification of data therefore becomes necessary and a minimal negative list should be prepared by the data generators and sources.

Data Sharing and Access policy shall apply to all non-classified data collected using public funds and held by various Ministries/Departments/Subordinate offices, etc. The Ministries/institutions/departments/sub ordinate offices etc generating and possessing data assets for sharing are hereafter referred to as organizations. Organizations managing the data shall decide the classification as to whether it is sensitive or non-sensitive from the security standpoint. Thereafter a negative list may be drawn of sensitive data. All data outside the negative list will be put on website for information.

For each data set, the format in which they are held, namely analog or digital, shall be identified. Effort should be to convert the analog data to digital form, if possible. This will enhance the portability of this data. Organizations shall prescribe procedures by which the access could be gained for specified and registered use. Organizations shall be free to fix a reasonable price for providing accessibility. All organizations managing data will classify the same as classified and non-classified.

Ministries/Departments of the Government of India shall be identified as a reference centre for mutual exchange for all organizations. The Ministries may evolve their own procedures for management and exchange the data managed by them. The administrative procedures should not be cumbersome so that the data transfer does not become time-consuming. The charges, if any to be levied should be decided by the concerned departments with a proviso that data costs do not become prohibitive for usage.

Data users will have to follow the guidelines governing the acceptable and registered use of the data. Guidelines will be formulated and placed on the website of the data managing organization.

### **3. CLIMATE CHANGE – A case study**

NAPCC has enunciated a total of eight missions. All these missions and other programmes in the context of the NAPCC would require extensive exchange and sharing of data among the agencies engaged in the implementation of mission-linked actions and other actions described in NAPCC. The concept of registered users could be developed and a suitable policy framework provided for data sharing. Under the current policy framework, ownership of data resides with the original generator. While sharing of data between one arm of the Government and another is permitted, the transfer of data by the receiving arm of the Government to a third arm is not allowed. Any value addition to primary data using service providers could be undertaken only within the premises of the owner department. The prevailing policy does not envisage value addition to primary data under public-private partnership using transaction or relationship models. It is important that the revised data sharing policy takes into account data types, technological, managerial, institutional, financial, legal and operational issues mentioned above.

A revenue sharing model to suit Data Sharing and Accessibility Policy without compromising to National security consideration is required. This would call for a consensus based approach for a collaborative excellence in developmental planning through innovative deployment of data generated with the use of public fund.

### **4. SALIENT FEATURES**

- Indian data available globally should also be available to the Indian citizens.
- The items excluded in the RTI Act will become part of the Negative List.
- The RTI Act only stipulates that Government provides information in a reactive manner and there is no provision for public information in proactive manner. There is a need to strike a balance between developmental needs and the security. Similarly, there should be a balance between Right to Information and Right to Privacy.
- RTI Act needs to be studied for the requirement of any explanatory notes.

- Negative List will be prepared by each Organization/Department/Ministry/taking into account the security, privacy, IPR etc.
- The Negative List has to be constantly reviewed so that it is realistic and is in tune with the technology.
- Policy should also not violate any existing laws such as IPR, Copy Right and proposed Privacy Law.

## **5. GLOBAL SCENARIO**

USA launched their open data through their [www.data.gov](http://www.data.gov) site which has now more than web site to provide data in a proactive manner as apart of open government initiative. When the U.S. government's data.gov site launched, critics pointed out that it was filled with relatively non-controversial data sets; plenty of USGS data but no Department of Justice or military data, for example. It has over U.S. site, Data.gov, has less than 1,000 data sets. A new website [www.data.gov.uk](http://www.data.gov.uk), dedicated to making non-personal data held by the U.K. government was launched and has more than three times as much data than the U.S. site offers. At launch, data.gov.uk has nearly 3,000 data sets available for developers to build mashups with. The The UK government has been a big supporter of innovation built on top of public data. The U.K.'s data site also includes 22 military data sets at launch. Still they are grappling with the interoperability and standardized formats. The U.K. site does include a prominent promotion of the Semantic Web focusing on the paradigm as the next step for the future of the web. More standardized, structured data is expected to be the direction that the program tries to get government agencies to move toward in the future. Canada is also actively pursuing 'Government 2.0' making data available at municipal, provincial and federal level.



