

# Towards an Expanded and Integrated Open Government Data Agenda for India

Sumandro Chattapadhyay  
HasGeek, Bangalore  
The Sarai Programme, CSDS, Delhi, India  
+91 78381 63651  
mail@ajantriks.net

## ABSTRACT

The paper analyses the recently approved and implemented government data policy (National Data Sharing and Accessibility Policy) in India, and identifies the possibility and need of expanding and integrating it within the larger national e-governance and information policy ecosystem. The study draws from an ongoing research project on the Indian government data policy and the roles played by data intermediary organisations. The paper argues that an expanded and integrated open government data agenda will address crucial shortcomings of the national e-governance initiatives in India, and strengthen democratic interactions between the state and citizens through electronic and other means.

## Categories and Subject Descriptors

K.4.1 [Computers and Society]: Public Policy Issues – *Ethics, regulation*; K.4.3 [Computers and Society]: Organizational Impacts – *Automation, computer-supported collaborative work, reengineering*

## General Terms

Design, Human Factors, Standardization

## Keywords

India, National Data Sharing and Accessibility Policy, National e-Governance Plan, Open Data, Open Government, Open Government Data, Right to Information..

## 1. INTRODUCTION

Yu and Robinson argue in a recent paper [1] that the concept of 'open government data' (henceforth, OGD) is a product of (and has simultaneously caused) a fusion of the concepts of 'open government' and 'open data'. While the former relates to ideals of accountable and transparent governance, the latter is connected with technical and legal standards for distributing structured data that allows for further re-usage and re-distribution of the data, including for commercial purposes. They note that an 'ambiguity' is created by an unfortunate blurring of the connection between the techno-legal availability of open government data and its socio-political impact. In a sympathetic critique of Yu and Robinson's argument, Peixoto emphasises the need to locate and analyse OGD agendas within local and national participatory governance mechanisms and civil society initiatives [2]. He

criticises the popular description of OGD as 'the new low-hanging fruit of good governance', and identifies the vibrancy of the public sphere context and the political agency of diverse members of the public sphere as crucial determinants of potential socio-political impacts of OGD.

The National Data Sharing and Accessibility Policy (henceforth, NDSAP) [3], prepared by the Department of Science and Technology during 2011-12, was approved and notified on 17<sup>th</sup> March 2012. Though the NDSAP document aspires to embrace both the economic value of OGD and its potential to enhance public accountability, it is located within the larger e-governance context in India that often takes a rather service-delivery-centric approach. Extending Peixoto's argument, presented above, regarding the need to analyse the potential of an OGD policy within its public/political context, this paper proposes that the OGD policy must not be studied as an autonomous initiative of the government but as tightly intertwined with the larger e-governance and public sector information policy situation. As part of an ongoing research work on the nature and functions of the ecology of data intermediary organisations in India, the paper analyses the OGD initiative in India within its national e-governance context, and identifies the need to expand the OGD agenda and integrate it with other policy instruments that facilitate informed and effective citizenship.

## 2. NATIONAL E-GOVERNANCE CONTEXT IN INDIA

The need for a national e-governance policy framework in India was first noted in the Report of the Working Group on Convergence and E-Governance for the Tenth Five Year Plan (2002-07) [4] produced by a Planning Commission committee in 2001. The document considered 'e-governance' as crucial driver in transforming governance from passive provision of information, goods and service to facilitating active and informed citizen participation. It also noted the challenge of re-engineering of government processes, moving from individual-driven to system-driven decision making processes, as a necessary condition for effective e-governance systems. The document identified not only technical requirements (e.g., large scale computerisation, local language enabled IT systems, standardisation, etc.) for a national-scale e-governance initiative, but also socio-behavioural ones (awareness development, knowledge networking, etc.). It also took note of a number of state-level e-governance initiatives already underway in India, including in Andhra Pradesh, Madhya Pradesh, Rajasthan and Tamil Nadu, as well as a national-level 'minimum agenda' of e-governance that was being pursued by NIC. The document concluded by emphasising the need for consolidation, standardisation and universalisation of the distributed and disconnected e-governance initiatives (as in 2001) through creation of a national 'master-plan of e-governance' and also establishment of a central agency to develop, inform and

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

ICEGOV2013, October 22-25, 2013, Seoul, Korea

Copyright 2013 ACM 978-1-4503-2456-4/13/10...\$15.00 ...\$15.00.

monitor the implementation of the 'master-plan'. National Institute for Smart Governance was established in 2002 and the Apex Committee for developing National e-Governance Plan (henceforth, NeGP) was constituted on 23rd August 2004. Completed and approved on 18th May 2006, the NeGP document envisioned national-scale implementation of e-governance initiatives through twenty seven Mission-Mode projects targeting various service delivery verticals, such as Income Tax, Central Excise and e-Panchayat, as well as development of Core and Support Infrastructures. For effective management of the NeGP, the National e-Governance Advisory Group was constituted in 2005 with representatives from central and state governments, public and private sector entities, trade association of Indian IT companies, and the academic community. The vision statement of NeGP declared its objective as to “[m]ake all Government services accessible to the common man in his locality, through common service delivery outlets, and ensure efficiency, transparency, and reliability of such services at affordable costs to realise the basic needs of the *common man*” [5].

This service-centric approach to the challenge of e-governance, perhaps, led the Apex Committee to develop not a 'national policy' but a 'national plan'. Instead of integrating electronic processes and devices into all aspects of everyday governance, the NeGP identified specific verticals of governmental activities, and proposed that complete digitisation of these activities should be carried out, driven by unambiguous political ownership at the highest level, well-defined standards and architecture, and service deliver partnerships between the public and private sectors. The document planned out a system of Core and Support Infrastructures to enable e-delivery of the services concerned. These infrastructural elements included e-Governance Service Delivery Gateways (standard-based interfacing, messaging and routing switch to ensure interoperability of communication and applications across government agencies and service-providers, at national and state levels, and also specifically for delivery of services over mobile devices), State Data Centres (state level hosting and delivery of e-services and application to be used by citizens, and also internally by government agencies), State Wide Area Network (converged backbone network for data, audio and video communication at the state level), and Common Service Centres (web-enabled rural kiosks to access high quality and affordable data, audio and video content and services, set up and managed through public-private partnerships to solve the 'last-mile'). These various aspects of NeGP propositions are now in different stages of implementation, and deployment of the infrastructural elements is expected to be mostly completed by end of this year.

To engage with the under-addressed issue of citizen participation facilitated by e-governance activities, the Department of Information and Technology released a draft Framework for Citizen Engagement in E-Governance [6] in August 2011. Along with suggestions such as organising consultation meetings with citizens at various stages of project planning and implementation, and participatory audit of ongoing projects, this framework identifies 'information sharing' as a critical strategy and prerequisite of citizen engagement. This provides an appropriate entry point to discuss the OGD initiative in India.

### **3. NATIONAL DATA SHARING AND ACCESSIBILITY POLICY**

A draft version of the NDSAP document was first made public on May 2011 and citizens' feedback was sought. Among cases of

submitted comments known to the author, the Bangalore-based Centre for Internet and Society shared a series of suggestions that were, however, not found reflected in the final version of the policy. The policy identifies that government produced data, collected and collated using public funds, if made publicly available in an organised, well-documented and timely manner, it can produce great socio-economic value and enable citizens in various ways. The document contextualises itself within both international citizens' rights declaration (United Nations Declaration on Environment and Development, principle 10) and national information disclosure law (Right to Information Act, 2005, section 4.2). However, it does not name any open government data initiatives undertaken by other governments, either as a point of reference or on a comparative note.

NDSAP is applicable to all ministries, departments, subordinate offices, organisations and autonomous bodies of the Government of India and mandates sharing of all 'shareable' 'non-sensitive' data through a common government data portal deployed and managed by the National Informatics Centre (henceforth, NIC), Ministry of Communications and Information Technology. While 'shareable' refers to data declared to be such by the government agency that created it, 'non-sensitive' refers to data sharing of which is not prohibited by any central government acts (such as unit level Census data). The policy clarifies that the shared data should be made available in both human-readable and machine-readable formats, should follow file and metadata standards as specified by NIC, and should be periodically updated. While it is exclusively targeted at central government agencies, there is an expectation that equivalent state-level policies will be passed in near future, thus allowing for open sharing of government data across different scales of the government. Although the policy upholds the principle of 'openness', in practice, the policy does not specifically require sharing of the 'shareable' 'non-sensitive' government data under 'open licenses' that allows for re-use and re-distribution of the data concerned. Hence, even though NDSAP satisfies various aspects of the 'open definition' [7] such as machine-readability of data and adoption of 'open standards', strictly speaking it may not be called an OGD policy. While, its contribution towards setting up an administrative-legal framework for opening up government data in India is substantial and fundamental, it also leaves a gap in the OGD policy.

Following the instructions of the policy document, NIC developed both an implementation guideline document [8] and the national data portal of India <data.gov.in>. The guideline document has already gone through multiple iterations and has incorporated various feedbacks, including from the Data Controllers (officers deputed to oversee and approve contribution of data to the portal, for each central government agency). This document details out the stages of the data contribution process, including the role and responsibilities of the Data Controllers, relevant metadata and file standards, and management of datasets after they have been published in the portal. The portal provides a unified catalog of datasets published by various central government agencies, including both data stored at the data portal itself and data stored in the server of the agency concerned. The portal allows users, both governmental and non-governmental, to browse the dataset catalog, view the metadata associated with each dataset, comment on and rank various aspects of the dataset, create basic visualisations by choosing variables from the dataset, download available datasets and submit request for those that are not available yet. The portal is powered by Open Government Platform <ogpl.gov.in>, an open source Drupal-based data and content management system, developed by NIC in collaboration

with the Office of Citizen Services and Innovative Technologies, General Services Administration, Government of United State of America. Leveraging and building on the best practices and features of <data.gov> and <india.gov.in>, the portal was first launched on 21<sup>st</sup> May 2012. Implementation of the NDSAP policy during the last year has seen a remarkable role played by the NDSAP Project Management Unit at NIC. Expanding its initial mandate of developing and managing the data portal, it has taken up roles as varied as enforcing adoption of 'open standards' for shared datasets, evangelising proactive sharing of government data across agencies through detailed consultation meetings, organising community outreach programmes to induce increased usage of the datasets available from the portal etc.

## **4. EXPANDING AND INTEGRATING THE OGD AGENDA**

An important shortcoming of the NDSAP document at its present form is that its mandate for proactive disclosure of government data is only applicable for central government agencies. It is, however, expected that the equivalent policies will soon be passed in the states. The discussion here assumes such a condition but the relevance of the discussion does not depend upon realisation of that condition. Further, as mentioned above, the NDSAP does not give a clear mandate for OGD but for proactive disclosure of government data. Thus, the initial challenges for the OGD agenda in India include development and adoption of 'open licenses' for sharing government data. The following discussion uses the term OGD while being aware of its limited applicability in the present policy situation in India.

### **4.1 Interoperability**

The supply-side of proactive disclosure of government data in India faces much difficulty due to insufficient cross-agency harmonisation of data operations. This issue is symptom of a more general challenge. National e-governance situation in India greatly suffers from lack of government-wide interoperability framework and national enterprise architecture. As NeGP does not propose implementation of standard frameworks to ensure cross-agency semantic and system interoperability, in reality the implementation of NDSAP may create a backwards push for the agencies to re-organise their data management practices. This provides a great opportunity to harmonise data management processes including both technical and semantic characteristics, across agencies. It is crucial to note that NDSAP conceptualises <data.gov.in> both as a portal for citizens to access government data, as well as for intra-governmental data sharing (often of 'sensitive' data). Thus, realisation of a more integrated OGD agenda may lead to re-engineering of agency-specific data management processes to align them to interoperability requirements for contribution to the data portal, and can produce highly cost-reducing and process-streamlining outcomes.

### **4.2 Decentralisation**

The argument for open sharing of government data is not merely about its instrumental value in making government information processes more efficient and enabling cross-usage of data among government departments. The draft Framework for Citizen Engagement in E-Governance [6] mentioned above, makes a similar mistake by locating 'citizen engagement' strictly within a government-controlled or institutionalised circuit of interactions. An important aspect of the OGD agenda is enabling discoverability and autonomous utilisation (commercial or non-commercial) of various fragments of government data, without it necessarily being part of a structured citizen engagement exercise.

Hence, the OGD agenda should actively facilitate and encourage public usage of government data by as disparate set of actors and stakeholders as possible. An important part of this goal is to decentralise the points of supply of open government of data. Like the OGD policies worldwide, NDSAP tends to push the point of supply of government data towards a central location. While this centralised catalog of all available open government data sets is highly useful, it needs to be integrated by individual and everyday experiences of accessing information. For example, the OGD web-infrastructure must be utilised to allow a person to walk in to a specific police station and see the neighbourhood-scale crime records data being shown on a screen and access the data in print/digital formats too. This surely is not an argument against centralised web-infrastructure for managing government data but for allowing administrative privileges to local creators, publishers, and re-users of government data (such as municipality authority and government schools) to make use of centralised web-infrastructure to serve open government data (and data-based services) at the most local level, especially to open up public participation and deepen accountability by sharing it with the agency's most immediate stakeholders. It must be emphasised here that such localisation of government data sharing and re-use should deploy open data management architectures (such as public data API) that would allow citizens' organisations to perform acts of sharing government data in case the relevant government agency is unable or unwilling to do so.

### **4.3 Integration with RTI Act**

The OGD agenda aspires to advance accountability of government agencies by subjecting them to independent public scrutiny driven by not only published information regarding the activities of the agency but also the data sets that informed and got produced by those activities. This aspect of the OGD agenda is well-understood and championed in the NDSAP document itself. Effective realisation of this potential of OGD, however, requires expansion of the proactive disclosure mandate and its integration with the reactive disclosure of public information agenda, that is the Right to Information Act, 2005 (henceforth, RTI Act) [9]. In this context, a recent office memo circulated by the Department of Personnel and Training is of great importance. This memo, circulated on 15<sup>th</sup> April 2013, specifies the guidelines for implementation of *suo motu* (that is, proactive) disclosure of information under section 4 of the RTI Act [10]. Section 4 of the RTI Act emphasised the need for proactive disclosure of government information, however, it did not specify how it was to be operationalised. The office memo fills up this gap by laying down guidelines for implementation of the section 4 mandate, as well as to ensure its compliance. It is very heartening that the guidelines make it obligatory for all government agencies to declare on the website the detailed list of all datasets managed by the agency concerned, with a mention of which of them are available publicly and which are not, as well as to publicly disclose all RTI requests received by the agency concerned and the respective responses given out. It goes on to mandate disclosure of multi-media information held by agencies (such as, video recordings of meetings and consultation sessions) and adoption of 'open standards' for sharing information and data.

As the implementation of the guidelines is expected to begin shortly, it is all the more important to design integrated workflows for informational practices across government agencies so as leverage the advantages of the three documents – RTI Act, NDSAP and the office memo about implementation of section 4. While, NDSAP covers proactive disclosure of datasets by

different government agencies, its compliance enforcing mechanism is much weaker compared to RTI Act. This reduces the former's potential for impact both in short and long run. On the other hand, RTI Act, as re-energised by the section 4 office memo, though makes it obligatory for government agencies to proactively disclose various 'shareable' and 'non-sensitive' government information in human- and machine-readable formats, it fails to specify the technical requirements of the content/data management system required for hosting and distributing the information. It neither mandates centralisation of web-infrastructure for digital disclosure of information and data under the RTI Act, thus multiplies the cost and effort for doing so at the level of each agencies. Clearly, a great potential exists for synergising the implementation strategies of RTI Act and NDSAP through an expanded and integrated OGD agenda.

## 5. CONCLUDING REMARKS

On the positive side, India is presently experiencing an emerging agenda for proactive disclosure of government data and information that is well-premised upon both the older (and relatively more driven by peoples' organisations) 'right to information' movement and the newer (and relatively more driven by technical and policy analyst communities) 'open data' advocacy. These forces, under an overall governance context that is transforming towards greater dependence on and sophistication of data-driven decision making, can create a potent impact for not only opening up government data for creation of various data products and services, but also for advancing a greater public monitoring of and participation in government activities. The existing research on the demand-side of the open data question in India, however, is severely lacking. It is not very clear how data and information published by Government of India, either as open data or not, flow through and is transformed and translated by various organisational nodes to ultimately effect and inform grass-root campaigns and focused advocacies. While, a dense and grounded ecology of intermediary organisations is necessary in converting and circulating government data and information to forms accessible and usable by citizens across socio-economic classes, any intervention towards that would require mapping the existing ecologies and its various capacities and challenges first. The ongoing research on data intermediary organisations in India, of which this report is a part, precisely attempts to do that.

## 6. ACKNOWLEDGEMENTS

The paper draws from an ongoing research study titled 'Opening Government Data by Mediation: Exploring the Roles, Practices and Strategies of Data Intermediary Organisations in India' undertaken by HasGeek Media LLP, Bangalore, India, with the author and Zainab Bawa as Principal Researchers. The study is

part of the 'Exploring the Emerging Impacts of Open Data in Developing Countries' research network managed by the World Wide Web Foundation and supported by International Development Research Centre, Canada. The author is grateful to Michael Gurstein, Tim Davies, Suchi Pande, Sunil Abraham, the members of the NDSAP-PMU team in National Informatics Centre, and especially to Zainab Bawa for insights and support. Comments by three anonymous reviewers were critically useful in giving the paper its final shape. Remaining mistakes and shortcomings are the author's only.

## 7. REFERENCES

- [1] Yu, Harlan & David G. Robinson. 2012. The New Ambiguity of 'Open Government'. *UCLA Law Review Discourse*. 59. Pp. 178-208.
- [2] Peixoto, Tiago. 2012. The Uncertain Relationship between Open Data and Accountability: A Response to Yu and Robinson's *The New Ambiguity of 'Open Government'*. *UCLA Law Review Discourse*. 60. Pp. 200-213.
- [3] Department of Science and Technology, Government of India. 2012. National Data Sharing and Accessibility Policy. The Gazette of India. March 17. Pp. 74-99.
- [4] Planning Commission, Government of India. 2001. Report of the Working Group on Convergence and E-Governance for the Tenth Five Year Plan (2002-2007). Tenth Five Year Plan Working Group Serial Number 83/2001.
- [5] Department of Information Technology, Government of India. 2011. Saarangh: A Compendium of Mission Mode Projects under NeGP.
- [6] Department of Information Technology, Government of India. 2011. Framework for Citizen Engagement in e-Governance (Draft).
- [7] Open Knowledge Foundation. Open Definition version 1.1. Accessible from <http://opendefinition.org/okd/>
- [8] National Informatics Centre, Ministry of Communication and Information Technology, Government of India. 2013. Implementation Guidelines for National Data Sharing and Accessibility Policy version 2.1.
- [9] Ministry of Law and Justice, Government of India. 2005. Right to Information Act.
- [10] Department of Personnel and Training, Government of India. 2013. Implementation of *Suo Motu* Disclosure under Section 4 of RTI Act, 2005 – Issue of Guidelines. Office Memo No. 1/6/2011-IR. April 15.