January 6, 2025

Enabling Data Flow for Effective Governance in India

By: Sunil Kumar

For stronger democracy and federalism, India needs data democratisation. Improved data flow between government levels and open access for all users will enhance policymaking, scrutiny, and evidence-based governance, especially at the local level.

India is on the cusp of a digital revolution. An optic fibre cable network is being laid to cover all gram panchayats and villages in the country. In two years, 5G has been rolled out in 80% of India's districts and the country expects to begin rolling out an indigenously developed 6G telecom network by 2028. This is expected to be five times faster than the 5G network and is likely to revolutionise, among other things, health, banking, education, and entertainment services as more and more people begin to use it. Smart cities and smart villages are expected to soon become a reality.

This, however, would pre-suppose the presence of an efficient digital network, greater digital access, vastly improved digital literacy, and data analytics using the latest artificial intelligence (AI) tools. While "control" and potential "use" of data by "big business"—whether national or multinational companies—remains a major policy challenge before national governments, in the Indian context, data democratisation is a major challenge, especially from the citizen's point of view.

Here, I mean democratisation of data controlled by the government. In my view, without data democratisation, governance will remain poor, federal institutions like urban and rural local governments will be emaciated, and democracy will not take deep roots in the country.

Government control over data

In India, it is the union and state governments that control most of the citizen-related data. One obvious example is the Aadhaar database where critical demographic and biometric data on more than 1.4 billion citizens is controlled by the Unique Identification Authority of India (UIDAI), a government organisation. The database aims to deliver good governance and ensure efficient, transparent, and targeted delivery of subsidies, benefits, and services to the people on it.

> Under the Parivar Pehchan Patra scheme, all families are provided a unique eight-digit family ID. It is linked to the birth, death, and marriage databases and is automatically updated when such events occur in a family.

Two other examples are the Kutumba project of the government of Karnataka and the Parivar Pehchan Patra scheme of the Haryana government. The Kutumba project aims to automatically deliver benefits and services to residents in the state by checking their entitlement and eligibility using data available in the state government's various databases. Under the Parivar Pehchan Patra scheme, all families are provided a unique eight-digit family ID. It is linked to the birth, death, and marriage databases and is automatically updated when such events occur in a family. It also enables citizens to view their entitlements and claim them.

These clearly show that huge databases are controlled by the union and state governments. In other states too, the situation is no different.

The evolution of information technology over the last three decades has generated a huge amount of digital data, which has led to the setting up of cloud storage and computing facilities, and mammoth data centres with huge energy requirements. There has also been data centralisation at the union and state levels. Depending on the scheme under which data is collected, its ownership vests with the union or state government. For instance, if it is a centrally sponsored scheme (CSS), where the central share is 60% and the state share is 40%, the data is owned by the ministry or department concerned. The source or origin of data is immaterial.

Hence, we have the curious situation where the custodian of data refuses to share it with officers at the state/district/block levels or with elected representatives of local governments. More than 90% of the data of most portals of the government of India is generated by the states at the village or city levels.

Thus, while data originates from gram panchayats or town/city levels of a state, the state or local government officials do not have access or permission to use it unless they receive specific permission from the union ministry concerned. State government officials are

reduced to petitioning the union government, and the chances are that they will not receive the permission in time. So, the first important conclusion is that the "generators" of data are not the "owners" of data.

Data integrity

The second important point is that it is only at the lowest level, which is closest to the people, that issues relating to data integrity can be best resolved. It is only at the local government level—the gram sabha in rural areas and the area sabha in urban local governments—that there is better scope for detecting data discrepancies and resolving them.

> It is only at the level of local government that data can acquire "a name and a face" and spur concrete, impactful action.

For instance, the dashboard of the Jal Jeevan Mission showed that a particular village in Maharashtra was being supplied piped water. However, the local elected representative could assert, based on his personal knowledge, that there was no supply of piped water in the village and that the information shown on the dashboard was "erroneous". Similar reports have been received from several states where gram sabhas refused to certify their status under the Jal Jeevan Mission because of deficiencies in water supply works and gram panchayats having to take over those assets for operation and maintenance.

This shows that issues relating to data integrity can be best resolved at the local levels. Departments have a tendency to engage in "advance reporting" so that their performance is not questioned by authorities at higher levels. It is only at the level of local government that data can acquire "a name and a face" and spur concrete, impactful action.

Let me refer to a news report published in *Hindustan* (Delhi edition, 6 October 2024), a Hindi daily, which highlighted the need to undertake data verification exercises at the gram sabha level. The report stated that between 1 April 2024 and 24 September 2024, 78,185 calves had been born in Uttar Pradesh according to data uploaded on the Bharat Pashudhan portal but that the animal husbandry department had administered Brucellosis vaccines to 314,602 newly born calves—roughly four times the number of calves born! Another news report mentioned that the husbands of 250,678 "widows" (who were receiving widow pensions) were drawing free rations in Uttar Pradesh!

If the data related to the birth of calves and their vaccination had been regularly shared with gram panchayats and presented before the gram sabha and discussed, the chances are that such data discrepancies would have been detected and resolved. In these cases, it seems like different arms of the same department were not sharing information. Further, no reports are generated and provided to officials and elected representatives of local governments that they can use to monitor or identify data discrepancies and take remedial action. It must be realised that there is no way in which such data can be analysed and acted upon even at the district level, not to speak of the state or central level.

Flawed digital data flow

The third important point is that the design architecture of almost all government portals is one way—data flows from the bottom to the top. Nothing comes down. So the third tier of government is sidelined in this digital data flow architecture. This design flaw, coupled with the tendency to control the access of even state officials to their own data, strikes at the root of federalism and is far removed from the concept of "cooperative federalism".

With a view to increasing transparency and promoting accountability in government, ostensibly large amounts of data are put in the public domain. But ordinary citizens find it next to impossible to navigate these portals, and most of the data is presented in a manner that makes it meaningless for them.

> It would make far more sense to villagers if 3D visualisation tools are used to show them where drinking water and sewer pipelines have been laid or not laid and what could be done to remove obstacles.

Even members of parliament (MPs) and members of legislative assemblies (MLAs) find it difficult to make sense of what is published on the dashboards of most government portals. In a way, this is tantamount to asking a right to information (RTI) petitioner to deposit Rs. 1 lakh as cost and wait six months for receiving 50,000 pages of data. It is drowning the citizen in a deluge of data.

If data is to become meaningful, it is important that the latest digital visualisation tools are used to undertake data analytics and present reports and information in a manner that citizens can easily understand and act upon. In meetings of gram sabhas/ward sabhas, where every voter is a member, it is critical that effective data visualisation tools are used to explain complex public policy issues in a language and medium that members can grasp.

It would make far more sense to villagers if 3D visualisation tools are used to show them where drinking water and sewer pipelines have been laid or not laid and what could be done to remove obstacles. This could also be a very effective tool in gaining public support for undertaking "anti-encroachment" drives in rural or urban areas. With effective data visualisation tools, evidence-based decision-making would get a boost and citizen participation in public policy formulation would increase manifold. This would also strengthen democracy at the grassroots because the administration would become more accountable to citizens.

Co-ownership of data

Finally, just as national governments are seized with the larger issue of "data sovereignty" and the relationship of national governments vis-à-vis big technology firms such as Google and Microsoft, in the national context it would be meaningful to examine whether state and local governments, the other two tiers of the federal structure, enjoy the same powers that national governments are seeking in the international context. All data is generated in the cities and villages of India and even the poorest Indian is contributing to the immense store of digital data, as proved by the number of Unified Payments Interface (UPI) transactions taking place every day in all nooks and corners of India.

> It may be meaningful to make local governments the "owners" of data generated from their area and grant them the right to study, analyse, and use such data to improve their functioning.

So the central issue here is whether the citizen has "ownership" of his own data or whether in the Indian context it is owned by all three tiers of government in their fiduciary capacity or just by the union government alone? We have seen how Indian citizens hand over their personal data to anyone who asks for it, more so if the person occupies a position of authority. In such a scenario, it may be meaningful to make local governments the "owners" of data generated from their area and grant them the right to study, analyse, and use such data to improve their functioning and thereby the quality of life of their residents. They also have to be legally responsible for ensuring data privacy.

A related issue, especially in the context of Indian democracy and ensuring a level playing field to both the ruling party and the opposition, is the access of political parties to the personal data of citizens, especially those of the "beneficiaries" of government schemes. At the moment, ruling parties have no qualms in using their power to gain access to the personal data of citizens (name, age, mobile number) who happen to be beneficiaries of government schemes and then using that to bombard them with propaganda at the time of elections. This flies in the face of the idea of "fair play" in democratic elections, besides eroding privacy, undermining individual autonomy, and even facilitating the emergence of a surveillance state.

Such wanton misuse of digital data related to citizens can be prevented if all three tiers of government are legally made responsible for ensuring data privacy. Stringent punishment should be prescribed if any government authority is found compromising data privacy.

It is said that digital data is the "new gold", and at some stage, sooner rather than later, data monetisation will happen. In such a scenario, it would be appropriate if all three tiers of government are legally recognised as owners of digital data, their shares are determined, and the proceeds of data monetisation are shared with them.

This should apply equally to government as well as private entities. Personal data cannot and ought not to be sold in the market or the "dark web". If local governments begin to reap some benefits from data monetisation (now largely restricted to private companies), they will also begin to address the issue of data integrity with greater seriousness.

Improved data analytics

It is equally important that universities and researchers are given access to the raw data now controlled by the union and state governments so that their research and data analytics not only enrich science but also help in improving governance, especially at the third tier. Unless and until governance improves at the local government level, there is no way in which we can even dream of becoming a developed nation.

If citizens are to have a role in formulating public policy and monitoring its implementation, universities, academic institutions, and civil society organisations will need to consciously partner with local governments, both rural and urban, in the country. They will need to interpret the data and present it to more than 3.2 million elected local government representatives and functionaries as well as citizens in a language that they can easily understand. It is only then that evidence-based decision-making will become a reality.

> Making data available to all users for clarifying, questioning, providing evidence or just building a factual base for policy making would immensely strengthen democracy and federalism.

It is also time to give the gram panchayats and other local governments the power to directly engage the services of consultants/experts who can undertake this exercise for them. This should be an admissible item of expenditure even for funds devolving to the local governments based on recommendations of the central finance commission or the state finance commission.

To conclude, at this point of time, it is critical that effective steps are taken by the government for data democratisation. Improving the digital data flow system between the union, state and local governments and making data available to all users for clarifying, questioning, providing evidence or just building a factual base for policy making would immensely strengthen democracy and federalism. It would also improve governance in the country at all levels, especially at the local government level. This would call for addressing issues related to domain, devolution and distribution of data in appropriate legislation and protocols.

Ensuring digital access and digital literacy also become extremely important in the Indian situation. It must be remembered that only by giving a name and face to data and its authentication by democratic public institutions such as gram/ward sabhas in gram panchayats and in area sabhas in urban local governments, can data integrity be ensured and the public policy-making process become meaningful in India.

Sunil Kumar is a visiting faculty at the Gokhale Institute of Politics and Economics, Pune, and a former civil servant. The views expressed are personal.