TANDEM RESEARCH

Putting the Cart before the Horse?

Response to Strategy for National Open Digital Ecosystems (NODEs)

Tandem Research / May 2020

The Ministry of Electronics and Information Technology (MEITY) put out a white paper seeking comments on National Open Digital Ecosystems (NODEs). The paper argues for the need to fundamentally rethink how public services can be delivered through the use of 'open, shared and modular digital platforms that seamlessly cuts across departmental silos'. In order to fulfil this ambition, the government has proposed a strategy for a public systems infrastructure through the creation of a national open digital ecosystem. According to the white paper, a NODE has been defined as 'open and secure delivery platforms, anchored by transparent governance mechanisms, which enable a community of partners to unlock innovative solutions, to transform societal outcomes'.

Tandem Research is an interdisciplinary research collective that generates policy insights at the interface of technology, society, and sustainability. We believe that evidence-based policy, supported by broad-based public engagement, must steer technology and sustainability trajectories in India. Our work seeks to ensure that no one gets left behind in the technology transitions shaping India's future. We appreciate the opportunity to submit our suggestions to the NODE consultation paper. Below we present a summary of our key points, followed by responses to the questions posed for consultation.

Summary

While national open digital ecosystems (NODEs) can enable citizen-centric governance and fuel innovation, several prerequisite conditions are not in place yet. These include varying levels of digitalisation, underdeveloped institutions for data governance, weak privacy and security infrastructure, and the absence of public deliberation around the purpose and governance of NODEs. Without addressing these issues, NODEs could amount to a case of putting the cart before the horse, with significant risks in terms of privacy, security, discrimination, and the unequal distribution of technology gains.

Open source and open standards can enable greater accountability and community participation. Open standards will enable a more vibrant and competitive marketplace, but only if the setting of these standards is also done in an open and transparent manner, and anchored in democratically accountable institutions. The question of open data needs to be publicly deliberated, to establish what constitutes a fair exchange of data, and how this might vary across actors. However, such openness will have to be weighed against considerations of data security and privacy. The example often cited in the consultation paper is of Estonia. Estonia, however, has chosen to prioritise privacy and security, by first, clarifying that the data belongs to the individual; second, limiting the openness of data; and third, limiting the use of data analytics.[1]

Governance frameworks will also need to be established for individual nodes, alongside a common governance framework that establishes a set of overarching principles and values, sets standards, and enables impact assessment and coordination across the nodes. We propose a set of key parameters and issues that must be considered for the governance of NODEs, highlighting democratic accountability as the cornerstone for such frameworks. Moreover, regulatory institutions need to build trust and capacity for safe, inclusive, and rights protecting NODEs.

There is also no direct connection between citizen-centric governance and spurring innovation, and these objectives can even pull in opposite directions. Ideally, the former should serve the latter, rather than both being positioned as equivalent objectives. Further, there is confusion between openness as a principle of governance and a technological standard - the former is primarily about transparency and accountability, and can even be achieved through non-technological instruments, As a technological feature, the promise is to enable open governance, but this rests on a number of assumptions and linkages that have not been specified. If open digital systems are to deliver citizen-centric governance, their design and governance must be anchored around active community participation, as well as the systems and capacities that can enable such engagement.

---

[1] Margetts, H., & Naumann, A. (2017). Government as a platform: What can Estonia show the world. *Research paper, University of Oxford*. Retrieved from https://www.politics.ox.ac.uk/publications/government-as-a-platform-what-can-estonia-show-the-world.html

Finally, the White Paper would benefit from a clearer articulation of purpose - of how these systems will bring public benefit to citizens and the mechanisms and supporting infrastructure that would be needed to support this. A framework of public benefit that recognises the opportunities and risks across various members of the community creates conditions and processes for identifying what constitutes a fair exchange of data, and the trade-offs entailed, is needed. The processes underpinning the creation of these platforms or infrastructure must be transparent and open to public scrutiny, and not private capture; without this, the foundations for creating open and citizen-centric governance are already crucially compromised.

We recommend that considerable progress has to be made on the building blocks of NODEs, before endorsing or recommending their establishment as envisaged here. We suggest that an experimental and iterative approach be adopted, testing frameworks, standards, licences, for particular sectors, or cities, with a smaller user base to identify community needs, concerns, and risks.

Key Questions for Consultation

1. On the Guiding Principles for NODES

This section comments on the principles proposed in the consultation paper. However, we also believe that an additional set of principles are required that i) outline the overarching purpose of a NODE; ii) establish a framework for public benefit and the various trade-offs involved, and iii) identify a set of values that will help navigate between these trade-offs.

For example, Data for Public Benefit, UK, outlines three conditions for the effective use of public data and provides a framework to consider public benefit.[2]

Data collection and use must be :

- **Purposeful**, i.e., provide direct and tangible benefits to individuals and communities
- **Proportionate**, i.e., actively minimise the amount of data needed to be shared, with clear parameters against the risk of data being used for purposes other than which it was shared
-

[2] Scott, K., Burall, S., Perrin, N., Shelton, P., White, D., Irvine, G., & Grant, A. (2018). Data for Public Benefit: Balancing the risks and benefits of data sharing. Involve. Retrieved from https://www.involve.org.uk/sites/default/files/field/attachemnt/Data for Public Benefit Report_0.pdf

- **Responsible**, i.e., it must be a more efficient and effective way to add significant value to a decision making or policy implementation than other approaches, and must be able to deliver the intended outcomes.

1.1. Principles for the Design of the Platform

1.1.1. Be open and interoperable

The principle of openness is certainly a useful one, but the paper lacks a clear definition of what is meant by open.

For accountability and transparency, NODES should be open source. Open standards will enable a more vibrant and competitive marketplace, but only if the setting of these standards is also done in an open and transparent manner, and anchored in democratically accountable institutions. There is nothing intrinsic about these principles of openness which guard against monopolies of unfair value capture unless participation in the NODEs is made open to all actors, instead of only a select few.

Whether databases should be open or not would depend on the type of data, its potential uses, and the efficacy of data security and privacy measures. A one size fits all approach is not recommended. Public consultation is needed to determine what constitutes a fair exchange of data, specific to particular sectors.

The Estonia experience provides an instructive example. In Estonia, while open standards and open source code are considered important, databases are only open to those who it pertains to, because of security reasons.[3]

Barcelona is also experimenting with technological tools that can enable citizens to have control over what data is open and who should have access to it. DECODE technology will be integrated into council services, allowing citizens to set specific 'entitlements' to personal data, such as who may access the data and for what purpose. The city's Chief Data Officer looks after the City's OS and decides who can access what data. Some of the datasets are made publicly available under different degrees of openness via APIs.[4]

However, these technological forms of openness will not in themselves promote citizen-centric governance. For example, in the case of open data, merely opening access to government

---

[3] Margetts, H., & Naumann, A. (2017). Government as a platform: What can Estonia show the world? Working Paper funded by the European Social Fund. Oxford Internet Institute, University of Oxford. Retrieved from https://www.politics.ox.ac.uk/publications/government-as-a-platform-what-can-estonia-show-the-world.html

[4] Bass, T. Sutherland, E. Symons, T. (2018) Reclaiming the Smart City: Personal data, trust and the new commons. NESTA: London, UK. Retrieved from https://pdfs.semanticscholar.org/d3d1/de77e2232d9ec7dd2856661e61da738094e4.pdf

datasets does not ensure greater government transparency or accountability. This is more so in cases where the data opened remains unreliable, unverifiable or otherwise prone to error, especially in a context where no liability is assigned to data providers. Further, the decision as to which datasets will be made available seems to rest on the government alone. Thus, from the perspective of transparent, accountable and citizen-centric governance, to be truly open would be to involve citizens in decision making about the purpose and governance of NODEs.

1.1.2. Ensure Security & Privacy

The white paper does not adequately address the issue of how security and privacy will be maintained. This must be a critical part of the framework. It is also important to recognise that there is a tension between the principle of openness and privacy/ security.

To use the example of Estonia that is cited in the paper, Estonia has chosen to prioritise privacy and security, by first, clarifying that the data belongs to the individual; second, limiting the openness of data; and third, limiting the use of data analytics.[5]

India's draft Personal Data Protection (PDP) Bill requires data fiduciaries to prepare a privacy by design policy. However, the bill fails to identify mechanisms for the implementation of such a policy. In addition to the features of privacy by design mentioned in the report, we further recommend the following principles:[6]

Conducting risk-benefit analysis to inform the design and implementation of open data programs
Conducting analyses, before the creation of open data ecosystems will help those designing the ecosystem to be mindful of the potential risks to privacy, and allow them to work privacy-protecting mechanisms into the architecture/design of such ecosystems.

Considering privacy at each stage of the data lifecycle
When designing such ecosystems, it is prudent to not only focus on preventing harm at the stage when the data is released. The right to privacy can be (and often is) violated at different stages of the process – collection, maintenance, release, and deletion.

Developing operation structures and processes that codify privacy management
Data management processes need to be stringent and uniform across the board. Additionally, sensitisation of all personnel involved is of utmost importance, since they will be responsible

---

[5] Margetts, H., & Naumann, A. (2017). Government as a platform: What can Estonia show the world. *Research paper, University of Oxford*. Retrieved from https://www.politics.ox.ac.uk/publications/government-as-a-platform-what-can-estonia-show-the-world.html

[6] Aneja, U., D'Cunha, J., & Ghildiyal, H. (2020). *Comments on the Personal Data Protection Bill 2019*. Tandem Research. Retrieved from https://tandemresearch.org/assets/Comments-on-PDP-Tandem-Research-25.02.20.pdf.

for the management of data.

Emphasise public engagement and public priorities as essential aspects of data management programs
For each prospective NODE, there must be an accompanying rationale as to why the data was made open, the potential implications for privacy, and consequent measures taken to protect privacy.

1.1.3. Adopt an agile, data-driven development method

We recognise the value of an iterative and experimental approach, but caution against the use of a start-up culture mantra for designing the delivery of public services or large scale public interventions. A fail-fast culture necessarily has some losers, and typically disproportionately affects the already marginalised or vulnerable. We need rigorous and publicly available risk and impact assessments, tethered to existing constitutional rights and human rights frameworks. A fail-fast culture must consistently and effectively provide alternatives for people that are of equal quality and bring equal benefits. For an agile process to be effective, it is essential that these processes are informed by a wide source of expertise and capacity. It also assumes, or requires, a fairly literate and well informed population.

1.2. Principles for Transparent Governance

1.2.1. Define accountable institutions

The consultation paper should define its vision of an accountable institution, without which a comprehensive evaluation of the NODE proposal is untenable.

In defining this, two key questions must be considered - accountability to whom and accountability for what. The whom here must be defined here in terms of the broader public. The what should be defined not only in terms of the efficient delivery of a service, but ensuring the principles of equity, non-discrimination, privacy and safety are equal priorities. As NODEs are envisioned to make a significant impact on the delivery of many public and essential services, democratic accountability must be at the cornerstone of every NODE.

SPVs can bring efficiency, but also risk conflicts of interest and a lack of democratic accountability. Existing SPVs are also known for side-stepping existing governance mechanisms that include elected municipal corporations and local level utility providers.[7]

[7] Praharaj, S., Han, J. H., & Hawken, S. (2018). Towards the right model of smart city governance in India. International Journal of Sustainable Development and Planning, 13(02), 171–186. doi: 10.2495/sdp-v13-n2-171-186

Instead, we recommend that this institution must remain a government body that is democratically accountable, and ultimately liable, for the ways in which NODEs are developed and used. The composition of this body and its governing principles must be transparent and arrived at through a wide and inclusive consultation process. Governing institutions will be needed for each node, as well as an overarching body.

Accountability must be ensured every step of the way - through the transparency requirements from all government and private actors, which could be implemented both by law and by design. Ultimately, accountability mechanisms must be designed with the end-user in mind- to what extent do accountability and transparency mechanisms enhance the agency of the user and give them control over their data? In the case of Estonia, for example, peeping into an individual's data is not allowed, and the individual is immediately alerted in case someone accesses their data.[8]

1.2.2. Establish Rules of Engagement

The rules of engagement must create an open and even playing field, which enables not only participation by the private sector, but also by communities, civil society, and individuals. They must also seek to prevent any conflicts of interest, and establish accountability and grievance redressal mechanisms. The rules of engagement must allow for regular community engagement, to understand needs and priorities, and use those to establish or attract new service offerings on the NODE. It is imperative that these rules are decided in a transparent manner, through a consultative process, including not only private sectors, but also civil society and sectoral experts.

1.2.3. Establish Transparent Data Governance

Rather surprisingly, the white paper does not make mention of the draft Personal Data Bill nor ongoing conversations around non-personal data. Referencing existing legislation is critical, and the NODES, at bare minimum, must be compliant with the Bill. Further, with many NODES, the issue will largely surround non-personal data. Currently, there is no consensus or clarity on how this will be regulated. We also maintain, similar to our response to the PDP, that a clear distinction cannot be made between personal data and non-personal data. With the continuously evolving concept of personal data in combination with burgeoning forms of data analytics, the 'datafication' of reality allows for any data to be conceived as personal.

Importantly, the PDP doesn't cover automated decision making and inferential decision making, which is likely to be a key aim and output of the NODEs. This is particularly significant owing to the platform's reliance on analytics, as is highlighted in the principle of vibrant

[8] Margetts, H., & Naumann, A. (2017). Government as a platform: What can Estonia show the world? Working Paper funded by the European Social Fund. Oxford Internet Institute, University of Oxford. Retrieved from https://www.politics.ox.ac.uk/publications/government-as-a-platform-what-can-estonia-show-the-world.html

community. The ecosystem is bound to generate copious amounts of data albeit not all sensitive; however, when datasets are combined, inferences drawn can lead to discrimination and even social sorting.

Estonia provides an instructive example, different from what is imagined in the NODE paper. The key enabler in Estonia is that data ownership has been clearly established as lying with the individual. Further, there are very clear rules around data being shared across organisational boundaries, in some sense, sacrificing the potential for big data analytics. The data registries are separated by organisational boundaries to prevent non-transparent linking of data sources. The entire system is designed to make analytics subject to scrutiny and certain barriers.

1.2.4. Ensure the Right Capabilities

Public sector officials must be trained and sensitised to the risks involved in open systems; equally, the specific and different risks that might arise with different types of openness must be understood.

San Francisco, for example, has developed an Open Data Release Toolkit to help municipal officials assess the utility and value of publishing a dataset against potential risks to individual privacy. The toolkit provides leaders with a clear, actionable process for minimising risks, allowing the city to use and release data in a more responsible, privacy-preserving way.[9]

To ensure that service delivery through nodes remains uniform and equally accessible to all users, both infrastructural readiness and technological expertise must be reinforced at regional and sectoral levels. Clear assessments of infrastructural readiness and availability of technological expertise is needed prior to the design and implementation of nodes.

1.2.5. Adopt a suitable financing model

Financing for the NODE could be either through a fee/subscription based model / for actors using the NODE to develop a service solution. The user fee could be calibrated according to the actor and the expected value gain or profit. In the case of community or individual use, for example, to address a specific need, the fees could be minimal, whereas for a private sector leveraging the NODE for a solution at scale, for commercial purposes, the fee could be significantly higher. However, as noted earlier, the blanket commercialisation of data should not be advocated as a suitable financing model - this would depend on the type of data and the ways in which it is being used.

---

[9] Bass, T. Sutherland, E. Symons, T. (2018) Reclaiming the Smart City: Personal data, trust and the new commons. NESTA: London, UK. Retrieved from https://pdfs.semanticscholar.org/d3d1/de77e2232d9ec7dd2856661e61da738094e4.pdf

1.3.Principles for Vibrant Community

1.3.1. Ensure inclusiveness

The community must be thought of as users, but also as builders and contributors to NODEs. Ensuring inclusiveness in this respect will require establishing the rules of engagement in an open, transparent, and inclusive way, that enable a wide range of actors to develop service solutions for community or commercial use through a NODE.

Inclusivity goes beyond mere design/language features and needs to take into consideration that people might lack the access, skills and infrastructure to use these platforms. This could include a range of things, from women having limited access to the internet because of prevailing socio-cultural norms in many parts of the country to low levels of digital literacy to unreliable or unavailable internet connections.

1.3.2. Be analytics driven and learn continuously

Once again, we would urge a closer look at the Estonia model, which has very clear rules around data being shared across organisational boundaries. The Saudi Arabia example cited in the paper seems misplaced. A crucial concern that has been raised with TAQAT is that it creates a barrier for the hiring of women. This highlights continuing worries around how biased data collection or data sets can lead to discriminatory outcomes.

1.3.3. Enable Grievance Redressal

This principle is better placed in the governance section. Moreover, it must extend beyond providing remedies for specific instances of harm, to making provisions for people and communities to access essential services, even for those without access to a node. In other words, systems must be structured to ensure that there is no exclusion or discrimination, by design, for those not availing of the NODE to access essential services.

2. On Delivery Platforms

2.1. Biggest challenges in Gov-Tech 1.0/2.0 to a NODE approach

The levels of digitalisation vary across the country, across states, and across government departments. Examples relied upon in the paper have contexts that are wholly different. Estonia, for example, had computers in schools in the 1990s, and 92% of the 1.2 million

population was online by 2016.Their long tradition of digitalisation started all the way back in 2005.

Data quality is also uneven and standards vary. Terminology and nomenclature in datasets, vary across different government offices.

Further, frameworks and institutions for data governance are still under development, and the processes leading to their development have been fraught with political contestation. Recent instances have also demonstrated poor understanding of privacy and security concerns at various levels of government.

NODES require strong regulatory capacity, to both understand and regulate for the complexity of the issue at hand, and then enforce those directives.

2.2. Should all delivery platforms be 'open source' or are 'open APIs' and 'open standards' sufficient? Please elaborate with examples.

Delivery platforms should be open source for transparency and accountability, as well to enable collaboration and innovation.

Many of the cities that are placing responsible use of data in their smart city strategies make a conscious effort to promote the use of open-source technologies. This is driven by the ethical argument that publicly funded technologies delivering a public good should be transparent and open to public scrutiny. But, using open-source also makes collaboration easier. A problem for smart cities is that technology vendors often build proprietary data solutions, which lack interoperability and can become technological 'silos' that make the council reliant on private technology providers.

A relevant example is the City of Barcelona's Digital City Roadmap, which promotes public digital infrastructures based on free and open-source software, open standards and open formats. It is rolled out in line with an ethical data strategy, where privacy, transparency, collective rights to data and other citizens' fundamental rights are core values.[10]

3. On Governance

In addition to responding to the specific question below, we recommend the need for greater

[10] Bass, T. Sutherland, E. Symons, T. (2018) Reclaiming the Smart City: Personal data, trust and the new commons. NESTA: London, UK. Retrieved from https://pdfs.semanticscholar.org/d3d1/de77e2232d9ec7dd2856661e61da738094e4.pdf

clarity on the *objective* of governance.

There seems to be a two-pronged objective driving the creation of nodes: one, to enable a citizen-centric model of governance and second, to fuel innovation, and bolster India's digital economy. These objectives can pull governance architectures and requirements in different directions, requiring at times, a prioritisation between the two.

If the objective is citizen-centric service delivery, then the role of the government has to be beyond that of an enabling platform for builders and function as an enabler of citizen autonomy and agency. Citizen rights have to be at the core of the delivery platforms - in both design, and implementation.

3.1. Do NODEs across sectors require common governance frameworks and regulatory/ advisory institutions to uphold these? Or is it sufficient for each node to have an individual governance construct? If a common framework is required, please elaborate the relevant themes/ topics e.g. financing, procurement, data sharing.

We recommend the creation of both a common governance framework as well as individual frameworks.

A common framework would establish a framework for public benefit and a set of guiding norms including a set of basic technical standards and rules of engagement; a clear framework for individual and collective rights over data, with the establishment of a data controller or data protection authority, whose task would be set limits on the ways in which data can be used as well as consider implications across sectors; and establish platform ownership and accountability to lie with a democratically accountable government body.

Individual governance constructs would be similar, but establish more specific conditions, around participation, data use and sharing, procurement, and grievance redressal.

The following issues and parameters will need to be considered for the governance of NODEs:

• ***Decision Making and Planning:*** Who decides when a node is needed? How will the feasibility of a node in particular domain/sector be decided, by what parameters? For instance, in the case of smart cities project, the decision to award the smart cities grant was done via open competition and proposal seeking - post which SPVs were set up and funds awarded.

• ***Sectoral parameters of nodes:*** The degree of openness in each node should be decided in consideration of sectoral factors - for instance, a health node is likely to be less open than a land ownership & transportation node. To what degree a node should be open, should be decided on a case by case basis, through public consultations, and with recommendations from

experts.

• ***Procurement and Participation:*** What will be the eligibility criteria? In the absence of technological state capacity to build and implement a node from end-to-end, what would be the nature of agreement between government & private companies? Will private companies be required to bid on specific tenders? How open will the bidding process be? Often the bidding process for technological projects have been known to be extremely exclusionary - with many smaller start-ups, with solutions to offer missing out due to procurement rules.

• ***Standards and Security:*** There should be clear guidelines on the choice of software development methodologies, technological directions and infrastructural standards. Data must be held securely, with any breaches being made public immediately or within a specific window of time.

• ***Financing:*** How will each of the nodes be financed? Will there be a standardized model of financing? The decision to adopt certain financial models over others must be informed by an understanding of the tradeoffs and benefits associated with the model. For instance, as noted above, charging a fee to users for certain services will widen existing socio-economic inequities.

• ***Data governance:*** Data should be portable and easily shareable for communal use, whereby profitability is not the main driving force. Data should be available via a privacy-aware infrastructure that allows data to be accessible on consent-driven terms (e.g. tools allowing people to keep personal data private or share for specific purposes). Governance mechanisms should also ensure that there are mechanisms to respond to potential harms caused by data use, including clear accountability. Along with privacy measures and personal data protection mandated by the PDP bill, there must be a clear framework of data governance in the case of nodes. Any use of data must be purposeful and proportionate to the underlying objective of each node.

• ***Access and communication:*** There must be clear rules to outline the usability and access to data including limits to data access. There should be clear rules of engagement and adoption of clear channels of communication between different government institutions and departments, as well as private actors. Communications has to be made transparent, particularly since private companies do not fall under the ambit of the RTI Act, the rules of engagement must place clear and strict transparency obligations on all enterprises in the ecosystem.

• ***Audits and system checks:*** There should be regular audits of the use of resources and funds to determine the costs of inputs and outputs and streamline spending priorities. Data use must also be audited to check for the responsible use of data. Audit reports must be made publicly available and clear guidelines must be established in the case of wrongdoing or exploitative practices in data use.

- ***Citizen rights and user agency:*** Citizens must have rights over their own data, and have the ability to control decisions over who to grant access to data - at different levels to different actors. Citizens should have the right to revoke access to their data.

• ***Grievance redressal:*** Efficient grievance redressal mechanisms should be in place - grievances should be resolved or addressed within limited timeframes and regulatory mechanisms should be put in place to ensure compliance with grievance redress timelines and quality.

• ***Impact assessment:*** Periodic impact assessments must be carried out to ascertain levels of utilisation and value addition by nodes, and identify any gaps and shortcomings.

3.2. What are some potential risks that open digital ecosystems can leave citizens vulnerable to, for example, risks related to data privacy, exclusion, having agency over the use of their data etc.? What types of overarching guidelines and/or regulatory frameworks are required to help mitigate them?

Privacy

Recent research shows that databases can never be perfectly anonymous, and an increase in utility of data is met with a decrease in privacy. Up until a few decades ago, methods such as anonymization (also referred to as de-identification) worked well since datasets weren't exhaustive. Newer datasets, on the other hand, contain hundreds or thousands of different pieces of information about each individual. These different pieces of information make re-identification easy. Developments in big data analytics have also given rise to concerns about group privacy, something most privacy frameworks don't specifically address. Many algorithm-based analysis methods group individuals together, creating collectives, insights from which are used to make decisions for the entire collective.

Addressing this requires a multi-pronged approach. New technological and design solutions are needed, such as experiments in differential privacy and privacy by design. But these will not be enough. We also need to establish collective rights over the collection and use of data, as well as strong accountability frameworks that hold data users to account for the misuse of data. Further, models of data stewardship and data trusts can be experimented with in an iterative and consultative manner.

Data Security & Misuse

History is testament to the failure of open digital ecosystems, including the ones repeatedly mentioned in this paper, to adequately protect personal data. In 2018, Integrated Health Information Systems (IHIS), the technology agency responsible for Singapore's public healthcare sector (SingHealth), suffered a cyberattack and ended up compromising the data of

over 1.5 million patients. The consultation paper does not adequately consider this issue, nor the tensions with the principle of openness. Robust data security systems and architecture will be needed, along with broader awareness and training for governments and NODE community members on basic data security hygiene and best practices. More broadly, preventing misuse will require clear limits on the use of public data, including limitations on its commercialisation. It could also include data impact assessments. For example, during the communal riots in Delhi, reports indicated that the government's Vahan database was used to target individuals belonging to particular communities on the basis of their names.

Exclusion and Discrimination

The notion of 'data as a single source of truth' in itself is a deeply flawed idea, stemming from the dissociation of datasets, from deeply historical, social and cultural as well as political practices of data collection. Data through enumeration and measurement is always a selection from the total sum of all possible data available - and thus are inherently partial, selective. One key example of this process are recorded instances where slums and informal settlements in India often go unreported or underreported in urban data collection and reporting.

Further, there are also ample examples of biases in datasets, stemming from historical practices of social exclusion and discrimination. For instance, crime data often tends to be biased towards certain communities, as they are heavily represented in these datasets due to entrenched societal biases and discriminatory beliefs. Similarly, the societal attitudes towards the elderly or women in India, often get reflected in health datasets, which are grossly under-representative of these social groups.

Given that the paper places emphasis on the data-driven analytics for designing solutions and decision-making around public services, it is important to recognise that flawed, incomplete and biased datasets may be counterproductive to citizen welfare, especially for those sections of the population which fall within the digital data gaps and lack digital identities and traces.

Addressing this will require multiple interventions - from building high quality, consistent and standardised datasets, to adjusting algorithms to compensate for data biases. However, eliminating such bias all together is not possible. Thus, it is essential that people continue to have alternative options to NODE services, responsive grievance reversal mechanisms, limitations to the purposes of a NODE, and the possibilities for algorithmic audit.

Loss of Human Intermediation

The paper mentions reducing government touch points as one of the benefits of the NODE approach. However, it would also result in the disintermediation of human agents, who are necessary, to provide the gaps and safety nets, as well as compassion and understanding.

Privatisation of governance

While open data has significant value for improving accountability for citizens, it also has a

flip-side i.e. the marketisation of public services. The benefits for entrepreneurs and business are likely to be significant, as they are engaged not only in commercial activities, but also, replacing the role of the government in delivering services. It's not just that open data *can* create new markets, a substantial portion of the push for open data that is *explicitly seeking to create new markets as an alternative to providing government services.* Open government data programs can be used as a form of privatisation and deregulation: a deliberate attempt to create new markets in public service delivery, instead of providing government services.

4. On Community

4.1. What are effective means to mobilize the wider community and build a vibrant network of co-creators who can develop innovative solutions on top of open platforms? What can we learn from other platforms or sectors?

Creating such a community will require two key things. First, the rules of engagement should be inclusive and transparent, allowing a wide variety of actors, including private companies, communities, and individuals to participate. Barriers to entry should be carefully scrutinised and eliminated which are primarily, or disproportionally, serving the interests of a small number of actors in the ecosystem. Second, capacity building will be needed, not just to develop new products and services, but also manage some of the challenges and risks posed above. The community is also not formed equally, and thus, there needs to be due consideration of the huge differentials in capacities and power within the community. Enterprises, for example, will have the resources to get the most out of open data as they will be able to apply the full range of big data capabilities to it.

Open data is expected to empower citizens, but will only empower those that have the capacities to make the best use because of the digital divide. The digitization of land records in Bangalore, for example, had the direct effect of shifting power and wealth to those with the financial resources and skills to use this information in self–interested ways.This does not mean that this is an inevitable outcome, but that without efforts to level the playing field, the community is likely to be a select few actors with advanced capacities, increasing thereby social divides.

4.2. Are you aware of any end-user adoption and engagement models that platforms have successfully adopted e.g. feedback loops, crowdsourcing use cases, offline awareness and on-boarding campaigns?

Activities like 'citizen sensing' are being trialled across various cities. If instead of asking what

data could be sold by one business to another, we ask what people really need from data, then very different answers come to the fore. For instance, 'The Bristol Approach to Citizen Sensing', developed in partnership with a think- tank based in Barcelona called Ideas for Change, is a strategy that the UK has adopted to place people at the heart of innovation in its smart city strategy. It involved three months identifying 'hotspot conversations' with local residents, and local groups who might be interested to take part. This was followed by another three months framing the issues, inviting community groups with an interest in the issues to meet one another, and clearly defining the problems that affected them. This included identifying problems facing local people while also exploring how technology could help them address it, and demystifying words like 'data'.[11]

Amsterdam has also shown how high- level policy initiatives can gain broad support from a wide range of stakeholders across the city. For instance, TADA (an anagram of 'data') is a manifesto designed to encourage more responsible use of data in Amsterdam.[12]

Providing digital tools is an important step for governments to take, but without supporting efforts to build awareness and capacity in the population there is a risk people will not adopt them. There are multiple examples of this in the recent past, from digital participation exercises, which attract no participation, through to open data portals, which are never being used. Simply building something is no guarantee people will either want to use it or know how to. Like in Bristol, the project shows how it is possible to unlock more of the communal value of the data by bringing people together to decide how it is used. Although technology played an important part as an enabler, success depended heavily on community mobilisation and outreach by the council.

4.3. Are you aware of any innovative grievance redressal mechanisms/models that go beyond customer support helplines to augment accountability to citizens? If yes, please describe along with examples.

A key component of grievance redressal must be the availability of alternative offline solutions, so as to not disadvantage citizens in any way. Effective grievance redressal will also require publicly auditable systems, which would enable explanations about why and how certain decisions are made about them. The New York algorithmic accountability task force was the first policy proposal of its kind, in terms of, a city-based institution with a mandate to inspect the operation of automated decision systems used by local governments.[13] Their initial

---

[11] Bass, T. Sutherland, E. Symons, T. (2018) Reclaiming the Smart City: Personal data, trust and the new commons. NESTA: London, UK. Retrieved from https://pdfs.semanticscholar.org/d3d1/de77e2232d9ec7dd2856661e61da738094e4.pdf

[12] Ibid.

[13] Bass, T. Sutherland, E. Symons, T. (2018) Reclaiming the Smart City: Personal data, trust and the new commons. NESTA: London, UK. Retrieved from https://pdfs.semanticscholar.org/d3d1/de77e2232d9ec7dd2856661e61da738094e4.pdf

proposals were radical, requiring vendors of algorithmic systems (often including private sector providers) to open all of their source code for public inspection. The strict requirements proposed by the bill were perhaps one of the reasons it struggled to gain widespread support, and was eventually watered down. Nonetheless the initiative has set a precedent worldwide for cities to take bolder steps in the monitoring of automated decision-making systems across the public sector and by private service providers.

5. On Support

Tandem Research would appreciate being able to contribute or provide feedback for further development of the NODE strategy and strongly recommends that the process be consultative throughout. In particular, we would recommend activities around citizen sensing, fair exchange of data, and interdisciplinary collaboration between technologists and social scientists to address governance issues and related risks. We would be keen to participate in this process and provide relevant inputs.

Tandem Research
343 Coimavaddo Quitla, Aldona
Goa 403508, India

www.tandemresearch.org
hello@tandemresearch.org