Background Note
National Data Governance Framework[1]

Introduction

Global data is expected to grow from 33 zettabytes[2] in 2018 by more than five folds to 175 zettabytes in 2025, 49% of which will be stored in public cloud[3]. Another research estimated that by 2025, the number of Internet of Things (IoT) devices will hit 10 times the world population -- about 75 billion[4]. These trends couple with the propagation of 5G networks and other next-generation devices will also equip the global society with data-centric applications in artificial intelligence, blockchains, augmented and virtual reality, and further boost demand towards a digital society, thus generating zetta more bits and bytes worldwide.

The need for government data is nothing new. For decades, the way in how government data is gathered, secured, used and shared has been an area of high interest to governments and the academic fields of development and public administration[5]. Government data has always had a key role but the ways in which data is created and used have changed dramatically, bolstered by the revolution in data technologies and the proliferation of applications of small and big data, real-time data, geospatial data and among others. In addition, with the emphasis of data gaps especially disaggregated data in the 2030 Agenda for Sustainable Development, at both global and national levels, data issues in the public sector have become increasingly prevalent in terms of governmental and academic interest, as well as real-world applicability. Data is now woven into every sector and function of government, just like other essential factors of its *raison d'etre* such as hard assets and human resource, and much of its operational activity simply could not take place without it.

But today's data proliferation also brings along a suite of risks and challenges including security, privacy and ethical issues, as well as a general lack of data literacy and institutional capacities especially in developing countries, countries with transition economies and countries with special needs. The exponential increase of data in the public sector, along with both its potential and challenges, has in turn translated to demands on effective data governance and related institutions. To add to the complexity of this subject, other than being one of the largest source of data producers and consumers in many countries, if not the largest, government also owns the critical role of a regulator for data, both governmental and non-governmental.

---

[1] Note: Authored by Wai Min Kwok, UN DESA (Source: Chapter 6, 2020 UN E-Government Survey)

2 Note: One zettabyte equals roughly one trillion gigabytes.

3 IDC. (2020). Global DataSphere. [online] Available at: https://www.idc.com/getdoc.jsp?containerId=ID C_P38353 [Accessed 12 Feb. 2020].

4 Statista. (2020). IoT: number of connected devices worldwide 2012-2025 | Statista. [online] Available at: https://www.statista.com/statistics/471264/iot-number-of-connected-devices-worldwide/ [Accessed 12 Feb. 2020].

5 Henry, N. (1974). Knowledge Management: A New Concern for Public Administration. Public Administration Review, 34(3), p.189. doi:10.2307/974902.

UN Peace and Development Fund Project on Data Management and Data Governance

The above-mentioned context defines the scope of the UN Peace and Development Fund's (UN PDF) capacity development project on "Developing institutional capacities for digital data management and cooperation to advance progress toward the Sustainable Development Goals". The aim of the UN PDF 2030-Sub-Fund Project is to implement innovative, forward-looking and pro-active projects in support of the 2030 Agenda and the Sustainable Development Goals (SDGs).

This project seeks to address the existing challenges and gaps, focusing on enhancing the institutional capacities of countries to utilize, manage and govern data in a comprehensive, objective and evidence-based manner, through regional and global cooperation. To this end, the project will:

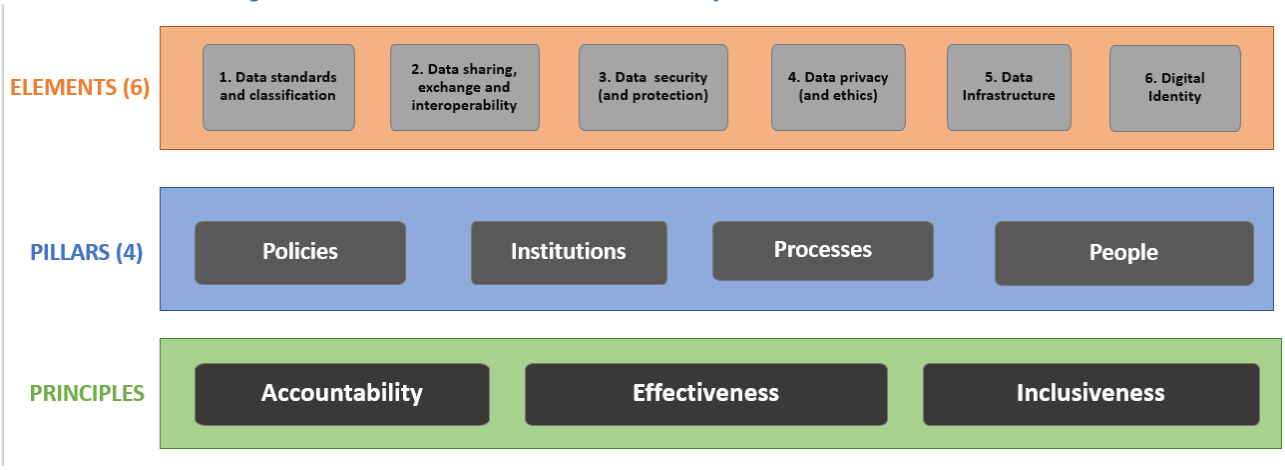
- Support African and Asian countries, especially those among LDCs and SIDS, to assess key data management and governance challenges and strengthen government officials' and stakeholders' knowledge of sound and secure data management, working closely with UN RCOs and UNCTs.
- Support national and regional initiatives in fostering an open, fair and non-discriminatory business environment for digital data cooperation through mutual benefit, win-win outcomes and common development.
- Support countries in developing institutional capacities for developing national digital data policies and strategies for ensuring data quality, access, security, privacy and usage, and for promoting data security through consultation, collaboration and shared benefits.
- Make available relevant legislative information and toolkits for advancing digital data cooperation agreements through case studies.

Establishing a national data governance framework

Establishing an effective national data governance framework is essential but difficult due to a series of paradoxes. There are various multi-faceted challenges in steering forward a data governance to drive data-centric e-government policies. For instance, in some legacies, data governance is still contained in IT or ICT governance which is not able to address the plethora of opportunities and challenges including data security and privacy. Another reason that data governance should not be embedded under ICT governance is that substantial amount of government data could be seen as "no person's land" -- where business users expect the responsibilities to come from ICT authorities, and yet the latter do not have enough context to fix data quality of other issues. The impasse could result in various issues affecting data integrity, data interoperability, security and privacy. In other instances, data governance could be implemented as an ad-hoc approach but not through a structured approach in supporting long-term sustainable development goals.

Data governance can, therefore, be seen as a homogeneous set of elements which assures formal management of data assets in an institution, with the fundamental aim to ensure that data are trustworthy and follow standardised processes[6]. Data governance also ensures that decisions based on available data do not place the institution or the public users at risk as a result of low quality, falsification of data, the use of obsolete data, or that could breach security or invade privacy. **Data governance can also be seen as a multiplex relationship of policies, institutions, people, processes and the effective use of technologies as catalysts or enablers.** More specifically, policies form the integral pillar to ensuring the legitimacy and institutionalisation of policy actions or inactions; institutions and people are another two integral pillars which comprises both governmental agencies as well as the general public; the ecosystem that includes processes such as data sharing and facilitating public participation. Figure 3 shows an illustrative national-level data governance framework for e-government.

Figure 1 Illustrative Data Governance Framework for E-Government



Source: Author (adapted[7])

Element One: Data standardization and classification

To ensure consistency of data treatment in the public sector, especially in integrated or whole-of-government approaches, a certain degree of data standardization is necessary. Enforcing this across specialized and autonomous government agencies across sectors, including those that operate in silos, is usually a central challenge of government.

Some countries have addressed issues related to standardization and classification through a leading ministry or an inter-ministerial commission or committee. For example, New Zealand's Statistics Office is the lead agency for government held data and it has guidelines on data standards and stewardship framework[8]. In instances where the government has adopted

6 IBM Data Governance Council (2008). Data will become an asset on the balance sheet and data governance a statutory requirement for companies over next four years.
7 Dizon D.M.H., Parcia, P.S., Kumar, R., Kwok, W.M., and Mukherjee M. (2016). 'Data governance in fostering policy coherence and collaboration for cleaning the River Ganga' in Poocharoen, O., Wasson R.J., and Wu X. (eds). Ganga Rejuvenation: Governance Challenges and Policy Options. Available at:https://www.worldscientific.com/worldscibooks/10.1142/9715#t=toc. [Accessed 12 Feb. 2020].
8 Stats.govt.nz. (2019). Data leadership. [online] Available at: https://www.stats.govt.nz/about-us/data-leadership [Accessed 14 Feb. 2020].

legislation, it is often part of the broader digital government strategy as it is the case of Colombia[9] and Estonia[10]. In the Republic of Korea, one of the leading countries as seen in UN E-Government Survey 2020, a series of guidelines focusing on data classification and standardization has been established, enforced and amended over the years to address the importance of standards and emerging trends, including its Guidelines for database standardization in public institutions[11].

Some countries have adopted **additional approaches that have sectoral approaches to data standardization or are more public-private-partnership oriented.** The European Multi Stakeholder Platform on ICT standardization was established to advise on matters related to the implementation of ICT standardization policies[12]. Japan has introduced public-private data utilization promotion law[13], as well as the Basic Law on Utilization of Public-Private Data that put an emphasis on promoting the use of public-private data through the construction of necessary infrastructure, the cooperation of business sector while taking the division of roles within the government into account. This serves to ensure cooperation among various entities that utilize public-private data, in order to maintain standards for information systems, ensure compatibility, as well as to ensure effective dissemination of public-private data.

Element Two: Sharing data; linking data; interoperability and data exchange platforms

One of the challenges with interoperability of government systems is a lack of cohesion in the way data are shared or managed, or the lack of it. At the horizontal level, one of the opportunities that governments should be keen to pursue, is combining data about an individual from several systems across agencies to gain a better overview of the individual, for example in providing e-services through a life-event approach[14] (e.g. in Singapore, Russia, Saudi Arabia). This will have transformative impact on a wide variety of elements, from the way they monitor the effects of specific initiatives to the way they deliver services to the public. Yet, to do so, the different departments have to establish tight collaboration in data collaboration which is often difficult in most bureaucracies.

Many countries have expressed interest and taken policy initiatives in embracing in the principles that underpin interoperability. In response, evaluating the frontiers of data interoperability policy and summing up guiding principles behind those progress and developments have becomes more conducive to the work of implementing interoperability for different agencies and stakeholders. Some have already specified technical requirements to establish or improve interoperability.

9SUIN-Juriscol MinJustice. (2017). DECRETO 1413 DE 2017. [online] Available at: http://www.suin-juriscol.gov.co/viewDocument.asp?id=30033063 [Accessed 17 Feb. 2020]. See also SUIN-Juriscol MinJustice. (2018). DECRETO 1008 DE 2018. [online] Available at: http://www.suin-juriscol.gov.co/viewDocument.asp?id=30035329 [Accessed 17 Feb. 2020].
10 Riigi Teataja. (2019). Riigi infosüsteemi haldussüsteem. [online] Available at: https://www.riigiteataja.ee/akt/129032016006?leiaKehtiv [Accessed 17 Feb. 2020].
11 국가법령정보센터. (2017). 공공기관의 데이터베이스 표준화 지침. [online] Available at: http://www.law.go.kr/admRulLsInfoP.do?admRulSeq=2100000122549 [Accessed 17 Feb. 2020].
12 European Commission. (n.d.). Internal Market, Industry, Entrepreneurship and SMEs. [online] Available at: https://ec.europa.eu/growth/industry/policy/ict-standardisation_en [Accessed 17 Feb. 2020].
13 電子政府の総合窓口. (2008). 平成二十八年法律第百三号 官民データ活用推進基本法. [online] Available at: https://elaws.e-gov.go.jp/search/elawsSearch/elaws_search/lsg0500/detail?lawId=428AC1000000103 [Accessed 17 Feb. 2020].
14 Note: Description of life-cycle approach

There are several options of sharing, linking or through a data exchange platform, that offers further digital services through linked data, data API, data service or data market[15]. Integration is key and connectivity is critical. Serving analytics use cases demands access to the data. It also requires the ability to integrate across multiple systems to serve people's needs, including user-centric policies such as once-only principles[16]. The pace at which new applications are built and the many technology options available mean a modern, open platform is critical as the connective glue that joins up analytical use cases.

Some governments have expanded from interoperability policies or guidelines to specific platforms such as payment. For instance, interoperability of data in Kenya is implemented through the National Payment System Act, 2011 provides at section 21 that a payment service provider shall use systems capable of becoming interoperable with other payment systems in the country and internationally[17].

The benefits of interagency and inter-levels (sub-national) government data sharing and exchange includes increased public-sector productivity, improved services, reduced data requests, evidence-based policymaking and integrating public services, and facilitating a whole-of-government or whole-of-society response to public needs and emergencies. In addition, this also facilitates innovative policies such as "once-only" data polices. For example, in China, in 2015, the General Office of the State Council of China issued the Notice on Simplifying and Optimizing the Public Service Process to Facilitate Grassroots to Start Businesses, requiring that "government agencies must not ask for proof materials from citizens unless they cannot get them through interagency sharing".

Element Three: Data Security

Almost every country has an incident of government data security breach, whether it is made public or not. Increasingly, there are more high-profile cases that have resulted in consequential economic or social loss.

It is estimated that the average cost of a data security breach will surpass US $150 million by 2020, with the global annual cost forecast to be US $2.1 trillion.[18] **Data breaches not only impair the effective functioning of institutions and impact the economic well-being of sectors like healthcare, but it also affects the safety and security of citizens.** There are also intangible social costs and contributes to the public trust deficits between people and their governments. For example, the healthcare sector frequently contains sensitive information that needs to be protected for privacy reasons and there are serious consequences at the aftermath of any hacking incident.

15 Notes on definitions and descriptions
16 Krimmer, R., Kalvet, T., Toots, M., Cepilovs, A. and Tambouris, E. (2017). Exploring and Demonstrating the Once-Only Principle. Proceedings of the 18th Annual International Conference on Digital Government Research - dg.o '17.
17 Government of Kenya. (2011). The National Payment System Act 2011. [online] Available at: https://www.centralbank.go.ke/images/docs/legislation/NATIONAL%20PAYMENT%20SYSTEM%20ACT%20(No%2039%20of%202011)%20(2).pdf [Accessed 13 Feb. 2020].
18 News.cuna.org. (2019). Data breach costs will soar to $2T: Juniper. [online] Available at: https://news.cuna.org/articles/105948-data-breach-costs-will-soar-to-2t-juniper [Accessed 31 Dec. 2019].

Along with the adoption or amendment of data policies and the request for more data and analytics capability in institutions, as shown in earlier sections, **the demand to enhance and enforce data protection and cybersecurity continues to rise.** The public is justifiably concerned about people's data being lost or stolen. Government has a statutory duty to protect the public's data, and as such it is vital that appropriate security measures are put in place to fulfil the responsibility to ensure online data security and protection as prerequisites to ensure the use of data to drive sustainable growth and maintain a healthy digital environment.

The advanced safeguard for governmental portals as a whole serves a critical role. In addition, adequate cybersecurity awareness, incident reporting framework and staff training are necessary in response to data breaches and cyberattacks[19]. Due to the level of uncertainty, cyber norms normally emerge before effective policy can be put in place[20]. Cybersecurity issues are of paramount concern to many countries, not just within countries but also among governments given the cross-boundary nature of the Internet.

Individuals also have an obligation to contribute to their own personal data security online. However, they can only be expected to act as responsible users if they understand what is at stake, are aware of the risks, know their rights, and have learned how to act[21]. Capacity development on cybersecurity and digital literacy in general should enable e-government users, including the vulnerable groups and minorities, to become more secure online and able to defend and demand their data security and safety[22].

Element Four: Data privacy and ethics

The rapid uptake of e-services and government data in the public sector have also brought about challenging privacy and ethics issues. Intuitively, governments will need to use large datasets including identifiable data to create good algorithmic models for policymaking. The call for greater accountability with data availability on a myriad of government programmes, ranging from education to housing to social protection, however, often conflicts with concerns about paternalistic approaches that government agencies may implement and trample on personal privacy.

The proliferation of data also brings about profiling and surveillance applications in the public sector, through which governments could survey and regulate the citizenry[23]. It is notable that government data use might not always be perceived to be in the public interest. Seeking consent of data use is also complicated as data ownership is not always clear, and it gets convoluted when data management is shared or transferred between agencies, making it at times impossible to trace accountability or attribution.

19 Davis, J. (2019). Massive SingHealth Data Breach Caused by Lack of Basic Security. [online] HealthITSecurity. Available at: https://healthitsecurity.com/news/massive-singhealth-data-breach-caused-by-lack-of-basic-security [Accessed 13 Feb. 2020].
20 Internet Governance Forum. (2019). Berlin IGF Messages. [online] Available at: https://www.intgovforum.org/multilingual/content/berlin-igf-messages [Accessed 13 Feb. 2020].
21 Internet Government Forum 2019. (2019). Security, Safety, Stability and Resilience. [online] Available at: https://www.intgovforum.org/multilingual/filedepot_download/9212/1804 [Accessed 13 Feb. 2020].
22 Internet Governance Forum. (2019). Berlin IGF Messages. [online] Available at: https://www.intgovforum.org/multilingual/content/berlin-igf-messages [Accessed 13 Feb. 2020].

Closely related to privacy issues are those of ethics that goes beyond the law. Ethics can be considered as a reflection of society's collective moral understanding[24]. The obscure challenge for governments is that some context of ethics could be codified in data policies or law, while some context is not. The judgements on the appropriate use of government data is then governed by a wider moral understanding. Ethics become more important when advances in technology are pushing the common understanding of the law to its limits, or when the law and policies are not in place. To add to the complexities, public perceptions across the social fabric are more diverse and shift over time and depend on a multitude of factors such as current global and national events and the media's (including social media) presentations of data privacy.

Along these lines, data-centric policies around digital government should always start with a clear policy or operational need and with a clear public benefit. Transparency and accountability measures are essential to make the case for the benefits of data initiatives and data sensitivities, and to avoid accusation of nefarious intentions[25]. There are emerging approaches to protecting privacy such as data triangulation[26], data minimization[27], data anonymization[28], differential privacy[29] and the use of synthetic data[30]; and in addressing issues such as disclosure of purpose, data use limitation and data retention.

In deploying digital government and data systems, it is important to put in place privacy statements and ethical frameworks. This a crucial start but it does not solve all the challenges as the nature of digitalisation is a constant change, creating new scenarios and risks routinely. Constant periodic review is needed in reviewing data approaches to review the applicability of data minimization, anonymization, algorithmic accountability and transparency. Both privacy laws and ethical implications require governments to understand public perception, including through e-participation, such as the nuanced notions and privacy needs among the vulnerable groups, so that governments can take appropriate corrective actions.

Element Five: Data Infrastructure

The effective implementation of a national data strategy needs to build on a data ecosystem -- one that includes data architecture, data cloud, visualization and analytics, engaging people and partnerships and foster data innovation. In considering the needs of data infrastructure, the sheer increase in the volume, variety and velocity of public data does not mean that there is a need to store all data. The obvious rationale is that storing unlimited or vast amounts of data without a purpose will become unsustainable in due course. Simply upgrading existing systems is often not sufficient to ensure that large volumes of data remain accessible and can be shared, used and analysed efficiently. A strategy will need to be in place to make informed choices about

24 Drew, C. (2016). Data science ethics in government. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 374(2083), p.20160119. [online] Available at: http://doi.org/10.1098/rsta.2016.0119 [Accessed 12 Feb. 2020].
25 Drew, C. (2016). Data science ethics in government. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 374(2083), p.20160119. [online] Available at: http://doi.org/10.1098/rsta.2016.0119 [Accessed 12 Feb. 2020].
26 Triangulation is described as an attempt to fully explain the richness and complexity of human behaviour by studying it from multiple standpoints (Cohen, Manion and Morrison, 2000)26. It has become a standard method when multiple sources of data have been used in a study.
27 Note: Definition on data minimization
28 Note: Definition on data anonymization
29 Note: That injects random data into a data set to protect individual privacy
30 Note: definition and application of synthetic data

what to keep and what to throw away. The very act of choosing forces a judgement to be made about what the purpose of the data might be, and it is possible to regulate data more intelligently based on that purpose, rather than trying to regulate the existence of the data itself (Boyd)[31].

Increasingly, more governments have taken the approach to transition to cloud infrastructures. In most cases, it has become clear that public data cloud may be a necessity when data size goes beyond a certain size, even though the transition from traditional relational database server to cloud server involves including technical, organizational and policy challenges. For examples, the European Union is developing cloud-based systems to facilitate access to data from its Copernicus environmental monitoring program[32], and state governments in India are turning to commercial cloud providers to facilitate citizen services[33]. In the Republic of Korea, the government has put in place Regulation on the development of cloud computing protection of its users[34]; efficient operation of administrative affairs[35] and on electronic government act[36]. One exception to considering government data clouds is when government data that relates to national security.

It is plausible that governments that have adopted a shared fabric among their data and analytics technologies, including through cloud service and on-premise infrastructure, shared tools, while also addressing related risks and challenges, will have an advantage as compared to those that retain a siloed approach. **Governments should also explore data partnerships, both as through public-private partnership as well as with multistakeholder partnership.**

Element Six: Digital identity

Establishing that someone is who she or he is – technically defined as "authentication", is an essential first step in the provision of any e-service to an individual[37]. **Digital identity, therefore, plays a central role in digital government development and data applicability, as it provides the basis through which data can be safely, securely and precisely shared within and between agencies to improve service and delivery**. For example, the foundation and success of Estonian e-government are widely held to rest upon its main pillar of electronic identity (eID) -- a chipped identity document that enables citizens to authenticate themselves electronically, access to both e-government and private services and digitally sign documents[38]. The 2020 Survey indicated that 125 out of 193 countries (65 per cent) have put in place in their government portals a digital identity authentication in accessing e-service.

31 Boyd, I. (2017). The stuff and nonsense of open data in government. Scientific Data, 4(1).
32 European Space Agency. (2017). Accessing Copernicus data made easier. [online] Available at: https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Accessing_Copernicus_data_made_easier [Accessed 12 Feb. 2020].
33 IANS (2018). State governments fast embracing AWS Cloud in India: Teresa Carlson - ET CIO. [online] Available at: https://cio.economictimes.indiatimes.com/news/cloud-computing/state-governments-fast-embracing-aws-cloud-in-india-teresa-carlson/65899758 [Accessed 12 Feb. 2020].
34 Korea Internet & Security Agency. Laws on the Internet and Information Security of Korea. [online] Available at: https://www.kisa.or.kr/eng/usefulreport/ictLaws.jsp [Accessed 12 Feb. 2020].
35 Law.go.kr. (2014). 국가법령정보센터 | 법령 > 본문 - 행정업무의 효율적 운영에 관한 규정. [online] Available at: http://www.law.go.kr/lsInfoP.do?lsiSeq=163382&efYd=20141119#0000 [Accessed 12 Feb. 2020].
36 Elaw.klri.re.kr. (2017). Statutes of the Republic of Korea. [online] Available at: https://elaw.klri.re.kr/eng_mobile/viewer.do?hseq=45844&type=part&key=4 [Accessed 12 Feb. 2020].
37 McLoughlin, I. and Wilson, R. (2013). Digital Government at Work: A Social Informatics Perspective. ISBN-10: 0199557721
38 e-Estonia. (2020). ID-card — e-Estonia. [online] Available at: https://e-estonia.com/solutions/e-identity/id-card/ [Accessed 12 Feb. 2020].

Conclusion

Maximizing the potential of government data as a strategic asset, or as "oil, gold, fuel, currency", and mitigating risks and challenges associated with data security and privacy (as observed in last sections) needs requisite capacities of public administrators to use and govern public data. This is, however, lacking in many public institutions especially in those countries with special needs, even though governments may be fully cognizant that making effective use of available data assets can have a positive impact on public service and public value delivery.

Data leadership is essential to drive the national data strategy, if it exists, or to implement the data governance framework. This often calls for an institutional review that could transform the way agencies, across sectors and levels, to effectively deploy government data as a strategic asset. Gaining support from the executive levels of governments for data initiatives is often framed as a common challenge especially among countries with low EDGI score. While this challenge relates to the difficulty of understanding and expressing the value that data governance and initiatives might be able to generate, achieving top-level support also has other objectives and consequences. For example, it may be easier to communicate the value of data to the rest of the organization, if senior decision-makers have understood it and helped frame the goals of data governance as related to the overall goals of the institution. As such, this challenge remains at the intersection of data exploration and exploitation: policymakers or leaders might not comprehend the value creating potential of exploiting data assets until they have seen successful examples of exploration[39]. The ability of an institution to conceive strategic direction for their data governance is often dependent on the top-level executives' capabilities for understanding the value creating potential of data.

When institutional reform for effective governance is not possible due to political or resource constraints, governments should not put aside the possibility of implementing incremental change[40]. **The first step is likely to be a push for some form of mandate for creating the institution and infrastructure needed for a data ecosystem or national data service, a central element as oversight body or a steering committee that would set out on leadership and performance indicators, review security and privacy measures, devise structured process and carry out strategic planning.** Pilot projects could be introduced to demonstrate quick wins and success on how data initiatives could be seen as a viable approach to addressing developmental challenges especially those that relate to the SDGs and national development goals.

---

39 Nielsen O.B., Persson J.S., and Madsen S. (2019). Why Governing Data Is Difficult: Findings from Danish Local Government. In: Elbanna A., Dwivedi Y., Bunker D., Wastell D. (eds) Smart Working, Living and Organising. TDIT 2018. IFIP Advances in Information and Communication Technology, vol 533. Springer, Cham.

40 Nielsen O.B., Persson J.S., Madsen S. (2019). Why Governing Data Is Difficult: Findings from Danish Local Government. In: Elbanna A., Dwivedi Y., Bunker D., Wastell D. (eds) Smart Working, Living and Organising. TDIT 2018. IFIP Advances in Information and Communication Technology, vol 533. Springer, Cham