aapti institute

Impact of the Non-Personal Data Governance Framework on the Indian Agricultural Sector

Impact of the Non-Personal Data Governance Framework on the Indian Agricultural Sector

Astha Kapoor, Sarada Mahesh
and Vinay Narayan

February 2022

Acknowledgements

This report was produced as part of Aapti's Data Economy Lab (DEL) initiative. In addition to contributions from the wider Aapti team, this report also draws from the expertise of numerous academic and industry experts and practitioners. We are grateful for their input during interviews and rigorous feedback sessions.

Editor: Yana Bey

Report design: Meher Rajpal | **Cover illustration:** Siddhartha Iyer

Aapti is a public research institute that works on the intersection of technology and society. It examines the ways in which people interact and negotiate with technology both offline and online.

This is a work of independent research, funded by
Microsoft.

Abbreviations

TRIPS	Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS)
API	Application programme interface (API)
AI	Artificial Intelligence (AI)
ACIPA	Australian Centre for Intellectual Property in Agriculture (ACIPA)
B2B	Business-to-Business (B2B)
B2F	Business-to-Farmer (B2F)
B2G	Business-to-Government (B2G)
CCI	Competition Commission of India (CCI)
DEPA	Data Empowerment and Protection Architecture Framework (DEPA)
EU	European Union (EU)
GDPR	European Union's General Data Protection Regulation (GDPR)
GODAN	Global Open Data for Agriculture and Nutrition (GODAN)
G2F	Government-to-Farmer (G2F)
HVD	High value datasets (HVD)
ICAR	Indian Council of Agricultural Research (ICAR)
IP	Intellectual property (IP)
MoU	Memorandum of Understanding (MoU)
MEITY	Ministry of Electronics and Information Technology (MEITY)
NZ	New Zealand (NZ)
NPD	Non-Personal Data (NPD)
NPDR	Non-Personal Data Governance Report (NPDR)
PDP Bill	Personal Data Protection Bill (PDP Bill)
PII	Personally Identifiable Information (PII)
PV Act	Protection of Plant Varieties and Farmers Act, 2001 (PV Act)
US	United States of America (US)

Key Definitions

Term	Definition
Data steward	It is a trusted intermediary who helps safeguard the rights of data principals while also unlocking data for public/societal good in a rights protective manner. Stewards allow data principals to exercise greater agential rights over their data and protect against harm and harness data value.
Farmer producer organization (FPO)	A legal entity formed by primary producers, viz. farmers, milk producers, fishermen, weavers, rural artisans, or craftsmen, which provides for sharing of profits/benefits among the members
High Value Datasets (HVD)	As proposed in the NPD report, a dataset may be considered of high-value when one or more of the following criteria are met: it contributes to transparency; its publication is subject to a legal obligation; it directly or indirectly relates to their public task; it realises a cost reduction; or it brings value to a specific target audience.
India Digital Ecosystem of Agriculture (IDEA)	IDEA is an evolving policy framework that aims to digitize the agricultural sector on a national level.
Non-personal data (NPD)	Any set of data which does not contain personally identifiable information, contains aggregated information generate by human activity, and data generated by IoT devices.
Non-personal data authority (NPDA)	The Authority proposed by the NPDR to design and enforce the regulatory framework on NPD in India.
Non-personal data Report (NPDR)	In 2019, the Government of India established a committee of experts to develop a regulatory framework. The committee of experts have released two reports for consultation, NPDR refers to the most recent public draft.
Personal Data Protection Bill (PDBP)	A Bill in India drafted to protect personally identifiable information. Currently in discussion, to be tabled in Parliament
Personally identifiable information (PII)	Any information that can be used to distinguish or trace an individual's identity, such as name, social security number, date and place of birth

CONTENTS

01. Executive Summary 8

02. Introduction 11

03. Approach & Methodology 16

04. Policy Analysis 20

4.1. Relevant data sharing policies 21

4.2. General critiques of India's data sharing policies 23

4.3. Agriculture sector-specific critiques of India's data sharing policies 24

05. Non Personal Data Framework 28

5.1. Overview of the NPDR 29

5.1.1. Definition and classification of NPD *30*

5.1.2. Formation of institutions for regulating NPD sharing *31*

5.1.3. Optimising value creation from high value datasets (HVD) *32*

5.2. Problems with the NPDR 33

5.3. Agriculture sector-specific problems of the NPDR 35

5.3.1. NPDR concerns for farmer communities *36*

5.3.2. NPDR concerns for businesses in the agriculture sector *37*

5.4. Role of Ministry of Agriculture and Farmers' Welfare 39

06. The Need for an Alternative Data Sharing System 43

6.1. Challenges of mandatory data sharing frameworks 44

6.2. Voluntary data sharing frameworks problems of the NPDR 47

6.2.1. International best practices *48*

6.2.2. Points of concern *50*

07. Recommendations 51

08. Conclusion 56

Annexures 58

Annexure A 59

Case Study 1 - Digital Green *59*

Case Study 2 - Jaljeevika *65*

Annexure B 73

Annexure C 74

01

Executive Summary

Sharing data for the purposes of social and economic development and innovation has gained importance in India in the recent past. This has been reflected as the core objective of policies such as the report by the Committee of Experts on Non-Personal Data Governance Framework (NPDR)[1]. The report, a revised version of which was released in January 2021, focuses on establishing standards for the collection and sharing of non-personal data (NPD) in India[2]. The recent report by the Joint Parliamentary Committee (JPC) recommends including non-personal data in the Data Protection Bill 2021 leaving open an option for a subsequent regulation on NPD. The recommendation heightens the urgency for policy-makers to adopt an evidence-based approach to regulating India's data economy. If finalized, India will be one of the first countries to have a "single national level regulation" that explicitly regulates the use and sharing of non-personal data horizontally across sectors[3].

The NPDR takes a structured approach to sharing of NPD based on the purpose of sharing. The NPDR notes that NPD sharing for business purposes already exists, and therefore makes no recommendations regarding the same, instead focusing on sharing of NPD for sovereign and public good purposes. The NPDR envisages the creation of High Value Datasets (HVD) to facilitate NPD sharing for public good purposes. HVDs can be created for several domains that can benefit the society at large, including agriculture.

In this context, this research report examined the possible impact of the NPDR on the agricultural sector with the aim to ground recommendations made by the Expert Committee into sectoral reality, and understand the impact that the NPDR is likely to have on the agriculture sector. The research included an examination of current Indian Government policies to regulate data sharing, with a focus on the agricultural sector, a review of current data sharing practices in the Indian agricultural system through a case study of two community level organisations and expert interviews, and international best practices from jurisdictions such as Australia, the European Union (EU), New Zealand (NZ) and the United States (US), whose agricultural sectors have adopted a bottom-up, voluntary framework for the data sharing process.

The NPDR defines a number of concepts which are broad and lacking in clarity. Such ambiguity (may) cause confusion for both farmer communities as well as agri businesses. For instance, the lack of clarity in the definition and functioning of data trustees within the NPDR is only adding more confusion at a grass roots level on how farmer communities will be able to protect their NPD. For small and medium businesses, there is confusion on whether their industry will meet the threshold of 'data drivenness', as they will be required to share data without any incentive to do the same, while also having to grapple with increased compliance costs. The lack of clarity

Image source: https://www.freepik.com/premium-photo/young-indian-farmer-standing-wheat-field_8179573.htm

[1] "Report Summary", PRS retrieved on September 8, 2021 from https://prsindia.org/policy/report-summaries/non-personal-data-governance-framework

[2] The Committee classifies non-personal data as any information that does not fall within the ambit of 'personal data' (as defined under the Personal Data Protection Bill) or data that is without any personally identifiable information.

[3] "Indian govt took first step to unlock value of non-personal data, must now bring in nuance", Chaturvedi et al., The Print (24 July 2020), retrieved on June 28, 2021 from https://theprint.in/opinion/indian-govt-took-first-step-to-unlock-value-of-non-personal-data-must-now-bring-in-nuance/467239/

also extends to other aspects including the guardrails that will be put in place to enable secure sharing of data, and the technical standards for interoperability and privacy protections. Finally, farmers' data is being collected without their meaningful consent, and data processing and sharing is being advocated in a legal ecosystem bereft of a data protection legislation; however, this is expected to be remedied soon with the introduction of a data protection Bill in Parliament.

Government policies on data sharing in the agriculture sector, a significantly complicated subject, possibly ignore wide and effective consultation with farmer groups and follow a top-down approach. Most policy documents are available only online and only in English – adding to mistrust and institutional issues being overlooked. The policies also may not address the low levels of digital literacy / access amongst farmers, meaning that farmers will not ultimately benefit from sharing their data given that they are unable to use digital devices required to access the benefits. The issue with the credibility and quality of agricultural datasets in India is well known, also need to be addressed and, the policies may not be accounting for this, thereby leading to unreliable datasets being shared for analysis. Unintended harms and exclusion of farmers from services arising from such analysis are not adequately addressed by the policies.

These issues remain largely unaddressed in new drafts of the data protection and sharing frameworks in India, including the NPDR. While the focus of policies has shifted to the interoperability of data in the agricultural sector, they are pivoted on the objective of economic development. As a consequence of this, farmers' interests remain de-prioritised. If left unaddressed, these institutional issues could stifle innovation in the economy and disincentivise stakeholders from participation.

The experience of international best practices shows that agricultural sectors in major economies have chosen to go with an ecosystem-enabled voluntary framework. Such a framework aims to promote altruistic sharing and innovation through the establishment of a trusted ecosystem, while reducing regulatory and capacity burdens across stakeholders. The working of the Ag Data Transparent[4] in the USA (which inspired the codes in Australia and New Zealand), and the EU Code of Conduct on agricultural data sharing show that investing in infrastructure, legal and technical building blocks that foster a trusted network for voluntary sharing while incentivizing stakeholders to participate in such a framework helps build policy that is reflective of varied needs and concerns across the board.

Image source: https://www.pexels.com/photo/white-sheep-on-farm-693776/

[4] Ag Data Transparent, https://www.agdatatransparent.com/

02

Introduction

Data sharing is an intricate part of the Indian digital economy. The flow of information is important for all stakeholders - businesses, consumers and the Government.

In the current data sharing landscape, businesses use this knowledge – be it in its raw or derived form – to evolve a better understanding of the demand side of the market. This is then used to improve the quality of services provided to consumers from whom they initially collected data. The sharing usually takes place either through contracts or equity arrangements[7]. These closed contractual relationships, however, are not always beneficial for all the parties involved, especially the data generators, who only receive value in the form of targeted goods and services and are unable to deploy their data more meaningfully. Issues of information asymmetry and powerplay come into the picture that may force less influential parties to consent to provisions they may not want to agree to[8].

The Government regularly collects data, at various levels, to promote public good[9]. However, researchers and businesses question the credibility of these datasets, owing to issues such as poor data quality, lack of consistent updating of information, and non-transparency (unclear methodologies of data collection)[10]. There are also concerns about potential surveillance risks as the Government amasses more and more data, and centralises information to track citizens[11].

The current data sharing system is flawed, haphazard, and often works against the interests of the data generators – citizens who are unable to bargain for their interests against the State or big businesses. To address some of these issues, especially as a way to rebalance power in the data economy with businesses, the Government responded with a plethora of regulations. However, none of these address vital questions of citizen empowerment, and have instead assumed the existence of smooth data flow relationships, while planning for large government led technological innovations and policy. At the core of all policy framing, there are two types of data that are shared – personal data (information that can be linked back to the individual who first shared it) and non-personal data (NPD) (any information other than personal data). In 2017, the Supreme Court of India[12] in the Puttaswamy judgment paved the way for the protection of personal data, urging the government to legislate on

Image source: https://www.freepik.com/free-photo/tractor-agricultural-machine-cultivating-field_11137387.htm

[7] Page 4 of the Non-Personal Data Governance Framework

[8] "Tech groups cannot be allowed to hide from scrutiny", Schaake M., Financial Times (January 15, 2020) retrieved on August 18, 2021 from https://www.ft.com/content/401ca73e-36b9-11ea-ac3c-f68c10993b04

[9] This is the first line under the heading 'Context setting – A case for regulating data' in the Non-Personal Data Governance Framework

[10] "Quality and credibility problems of govt data", Ghosh S.K., Business Standard (December 3, 2019) retrieved on August 18, 2021 from https://www.business-standard.com/article/opinion/quality-and-credibility-problems-of-govt-data-119120300029_1.html

[11] India's personal data privacy law triggers surveillance fears, Gopalakrishnan, M (November, 11, 2020) https://www.dw.com/en/indias-personal-data-privacy-law-triggers-surveillance-fears/a-55564949

[12] Justice K S Puttaswamy versus Union of India, WP (C) 494/2012

the issue. Consequently, the Personal Data Protection Bill[13] (PDP Bill) was drafted in 2018, and is currently being deliberated upon by the Joint Parliamentary Committee on the PDP Bill[14]. At the time of writing this report, the PDP Bill was still awaiting Parliamentary approval. The discussion on NPD is more recent. The Ministry of Electronics and Information Technology (MeitY) constituted an Expert Committee on NPD in September 2019[15].

This Committee released its report on NPD (NPDR) in 2020[16] and was cognizant of the need for a separate legislation for NPD. If passed, this will be the first set of regulations on the subject in India[17]. The NPDR is also the first policy document that prioritises the interests of communities, the originators of the raw NPD, by setting up institutions to protect them from any harm that may ensue when their information is put to use. It also mandates that 'data businesses' – companies with significant control over the demand and supply sides of the market – will share their intelligence with the government, who in turn will make it accessible to other stakeholders. This move is critical for encouraging the setting up of start-ups, supporting small businesses, and giving back to communities. The NPDR received a mixed response from the public. Efforts to set up a committee to look exclusively into the issue of NPD were appreciated, though some questioned the need for this new category of data all together. The decision to look at this data as public good, thereby prioritising the interests of communities, was also highlighted. Making data non-excludable challenges the powers of big businesses and paves the way for new companies to enter the market[18]. Criticism of the NPDR stems from the fact that its provisions were vague

and provided little visibility for businesses and communities to prepare for the changes to come. Responses to the Committee included the need to differentiate between personal and non-personal data, clarity in the roles of data trustees and data businesses, better understanding of group privacy (and a mechanism for addressing situations of conflict with individual privacy), and the need for a better explanation of the technology architecture involved in the sharing process[19]. However, the biggest criticism of the NPDR emanates from its suggestion to make data sharing mandatory for public good, which has the potential to encourage regulatory arbitrage and can disincentivise investments in data generating, and processing activities in the private sector[20].

The NPDR highlights several sectors that can be the site for NPD sharing for public good – retail and e-commerce, transport, smart cities,

Image source : https://www.pexels.com/photo/woman-carrying-bag-with-cane-in-field-6026441/
[13] Available at: http://164.100.47.4/BillsTexts/LSBillTexts/Asintroduced/373_2019_LS_Eng.pdf
[14] Details of the Committee can be accessed here: https://prsindia.org/parliamentary-committees/joint-committee-on-the-personal-data-protection-bill-2019
[15] Official notification announcing the Constitution of a Committee of Experts to deliberate on the data governance framework: https://www.meity.gov.in/writereaddata/files/constitution_of_committee_of_experts_to_deliberate_on_data_governance_framework.pdf
[16] Available at: https://ourgovdotin.files.wordpress.com/2020/07/kris-gopalakrishnan-committee-report-on-non-personal-data-governance-framework.pdf
[17] "Brief Analysis Of Non-Personal Data Governance Framework In India", Dalmia V., Mondaq (November 5, 2020) retrieved on August 10, 2021 from https://www.mondaq.com/india/data-protection/1000976/brief-analysis-of-non-personal-data-governance-framework-in-india
[18] "India needs to be more mindful in regulating non-personal data", Deccan Chronicle (August 8, 2020) retrieved on August 18, 2021 from https://www.deccanchronicle.com/technology/in-other-news/080820/tech-this-week-india-needs-to-be-more-mindful-in-regulating-non-pers.html
[19] Responses referred to: Dvara Research (https://www.dvara.com/blog/2020/10/06/our-response-to-the-report-by-the-committee-of-experts-on-non-personal-data-governance-framework/), iSPIRIT (https://pn.ispirt.in/ispirts-response-to-non-personal-data-governance-framework/) and GFMA (https://www.asifma.org/resource/gfma-response-to-meity-india-non-personal-data-framework-consultation/)
[20] Mandatory Sharing of Data May Encourage Regulatory Arbitrage, Sonkar, S & Jayaram J (January 21, 2021) MediaNama Retrieved from https://www.medianama.com/2021/01/223-mandatory-sharing-non-personal-data-regulatory-arbitrage/

healthcare, and agriculture. The application of the report's recommendations across sectors remains unclear, and needs to be examined more closely given the diversity in data access, and variance in quality, and ability of stakeholders across sectors. For instance, the contribution[21] of the agricultural sector to the country's GDP is the highest – it increased from 17.8% in 2019-20 to 19.9% in 2020-21. Despite this, the sector continues to see low levels of digitisation and digital literacy amongst its stakeholders, especially farmers who are core to the sector[22].

The current agricultural data sharing system is scattered and imbalanced, with only a few stakeholders getting access to data and thriving off of it[23]. NPD for the sector includes weather data, soil data, and fertiliser data which can be used to provide agricultural advisories to farmers. This can help improve the output while making best use of the available resources, which benefits farmers and businesses through improved quality and increased income. However, big agricultural businesses are able to leverage their infrastructural power and networking capabilities to enter into contracts with governments. Medium and small businesses and start-ups have little negotiating power in contracts, putting them at a competitive disadvantage. This "economic moat" created by big businesses is supported by the government, mostly by its ominous silence around the issue[24]. These sectoral peculiarities need to be examined further, especially as they interface with recommendations such as mandatory data sharing and the impact it can have on stakeholders, especially farmers.

In this context, this research report unpacks the impact of the NPDR on the agricultural sector and aims to ground recommendations made by the Expert Committee into sectoral reality. To this end, this research begins with an examination of the current policies introduced by the Government of India to regulate the digital economy. The policies are scrutinised in detail to highlight the important on-ground issues that they fail to address. A detailed overview of the NPDR is then provided, along

Image source: https://www.freepik.com/premium-photo/young-indian-farmer-plowing-field_8179477.htm

[21] Economic Survey 2020-2021, available at: https://prsindia.org/policy/report-summaries/economic-survey-2020-21

[22] "The digital dream: Upskilling India for the future", Mothkoor V. et. al., Ideas for India (March 23, 2021), retrieved on August 2, 2021 from https://www.ideasforindia.in/topics/governance/the-digital-dream-upskilling-india-for-the-future.html

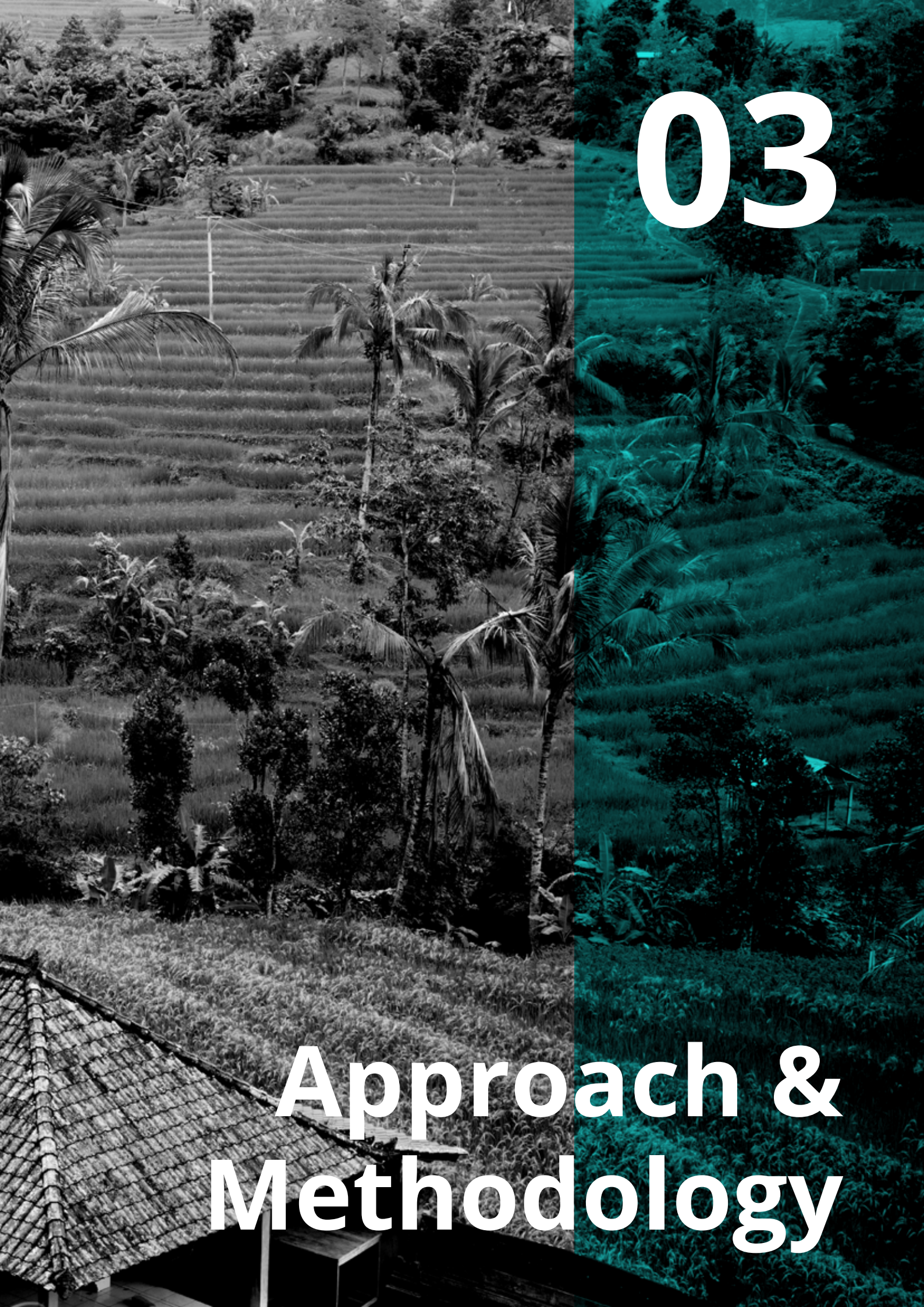
[23] "The global divide in data-driven farming", Mehrabi Z., et. al., Nature Sustainability (November 2, 2020) retrieved on August 18, 2021 from https://www.nature.com/articles/s41893-020-00631-0

[24] "How big business exploits small business", Stewart E., Vox (June 30, 2021) retrieved on July 30, 2021 from https://www.vox.com/the-goods/22550608/how-big-business-exploits-small-business

with the reasons for setting up the Expert Committee, the process set up for public consultation, and responses to the latest draft of the NPDR. Thereafter, current data sharing practices in the Indian agricultural system are analysed – this review encompasses loopholes in the system, businesses' and communities' contribution to the datasets and their reluctance to engage with the same for their work, and the legislative response to this issue. The report then discusses two types of data sharing frameworks – voluntary and mandatory – and highlights international best practices from jurisdictions such as Australia, the European Union (EU), New Zealand (NZ) and the United States (US), whose agricultural sectors have adopted a bottom-up, voluntary framework for the data sharing process. It brings forth the pre-existing issues that have to be resolved before implementing the provisions of the NPDR, which, left unaddressed, may actually widen the existing gap between stakeholders. The research also looks at community level approaches to data sharing, and how those are organised to benefit the farmers from the perspective of two non-profits, Digital Green and Jaljeevika.

The current recommendations of the NPDR are conceptual and high-level. Discussions about piloting the framework or creating more evidence on the ground have not been adequately addressed. Therefore, this research argues for sectoral import of these provisions, rather than imposition as an umbrella policy that may not address technical nuances of sector-specific stakeholders. We hope that this research is considered a starting point for a more nuanced, sector-specific discussion on the NPDR, and can serve as a launchpad to test, pilot and move towards more evidence based policy making on questions of data.

Image source: https://www.freepik.com/free-photo/paddy-harvest-golden-yellow-paddy-hand-farmer-carrying-paddy-hand-rice_14778442.htm#page=2&query=crops&position=47



03

Approach & Methodology

The conversation on the NPD and its application is still very conceptual, as a result, the literature on this question is limited.

This report is the first comprehensive study that uses a multi-pronged approach to analyse the impact of the NPD on a particular sector – it combines a theoretical understanding with practical, implementation based perspectives, through interactions with businesses that work with communities on the ground.

The NPDR is an umbrella policy that has been so far designed to broadly govern all NPD sharing for public purposes and does not take into account the needs of different sectors – levels of digitisation, capacity, potential for harm, etc. In Figure 1 below, it is clear that while retail and e-commerce are highly digitised sectors that will make up for a large share of the economy in the next few years, data sharing in agriculture is likely to create to more social/public good since it is a huge source of employment and economic value to the country and likely to be a focus area for NPD recommendations[25].

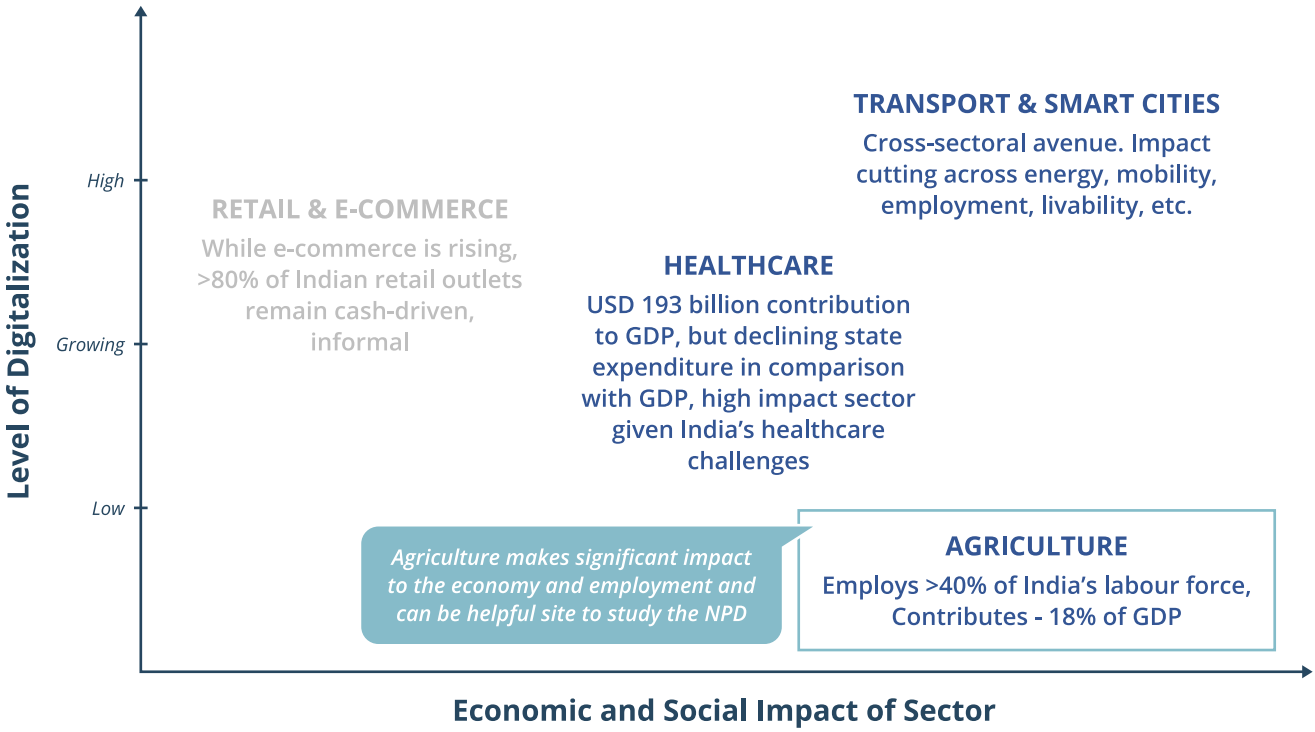


Figure 1: Sectors of interest for NPDR

Image source: https://www.freepik.com/free-photo/farmer-standing-rice-field-with-tablet_3738144.htm#page=1&query=agriculture&position=4

[25] "Agri share in GDP hit 20% after 17 years: Economic Survey", Kapil S., Down To Earth (January 29, 2021), retrieved June 28, 2021 from https://www.downtoearth.org.in/news/agriculture/agri-share-in-gdp-hit-20-after-17-years-economic-survey-75271

However, the agricultural sector, despite its potential, faces a twin problem of low levels of digital literacy and a morass of policies introduced by the government to digitise what has been an extremely traditional and informal sector. According to a study[26] by NITI Aayog, casual workers in the agricultural sector have the lowest levels of digital literacy, at 13%. The number lies at 24% for those who earn a regular salary working in the agricultural sector, and 26% for those who are self-employed in agriculture. These are extremely low numbers in comparison to the non-agricultural sector – (15%15% for casual labour, 53% for regular salary earners and 32% for the self-employed). This gap in knowledge leads to information asymmetry[27] – with limited information available in their hands, farmers will not be unable to negotiate benefits effectively with large agricultural businesses. These statistics are only symptomatic of the broader issues around digital preparedness in the agricultural sector, which raises questions about whether it is possible to implement the NPD framework, as it stands right now. These contradictions make the agriculture sector an interesting site for examining the impact of the NPDR recommendations.

To understand some of these sectoral complexities, the analysis in this report occurs at three levels. First, a top-down policy perspective to analyse the ongoing conversations on data sharing in India, which include the PDP Bill, NPDR, and other agriculture specific efforts such as Agristack. Policy analysis also includes a scan of international best practices for data sharing in agriculture, identifying both procedural and content related insights. The research delves into farmer-group led provisions such as in the EU, NZ and US and demonstrates how voluntary frameworks led by impacted stakeholders are the most effective in institutionalising data sharing (see Annexe B for detailed list of policies reviewed). Second, a bottom up, community led perspective to bring forth the views of impacted groups such as field-based farmer focused organisations. To examine the impact of the NPDR on the communities, this research studies the activities of Digital Green and Jaljeevika, two India based farmer organisations that use technology and data to create impact on the ground. Beyond ease of access, these two organisations represent two ends of the spectrum on tech-enabled decision-making and action – Digital Green is technologically advanced and uses technological and data-innovations in their day-to-day activities; Jal Jeevika has started to venture into ideas of deploying data for community benefit only recently. In both cases, the objective of the research was to understand how voluntary sharing occurs between farmers and organisations and how value of data is understood and enhanced. Finally, this research takes into account the perspectives of small and large agricultural businesses that form the backbone of India's farm sector to understand the implications of the NPDR. This analysis also focuses on the specific impact that mandatory data sharing can have on businesses and incentives tied to data collection, processing, and deriving insights.

Image source: https://www.freepik.com/premium-photo/indian-farmer-working-traditional-way-with-bull-his-farm-indian-farming-scene_9170688.htm

[26] "The digital dream: Upskilling India for the future", Mothkoor V. et. al., Ideas for India (March 23, 2021), retrieved on August 2, 2021 from https://www.ideasforindia.in/topics/governance/the-digital-dream-upskilling-india-for-the-future.html

[27] "Farm Organisations Flag Privacy Concerns, Seek Withdrawal of Agristack by Centre", Mohan Ram G., The Wire (July 10, 2021) retrieved on August 2, 2021 from https://thewire.in/agriculture/india-farmers-agristack-concerns-technology-agriculture

In writing this report, we conducted rigorous desk research, reviewing policy frameworks in India and globally to understand data sharing structures in general, and specifically to agriculture. This was supplemented with interviews with academics, researchers, policymakers, and data scientists from India and across the globe. The focus of these conversations was on past sectoral policies, and the advantages and disadvantages of the different data sharing frameworks while also trying to provide possible alternative solutions, all of which form a large part of our report (see Annexe B for a full list of interviewees).

To fully unpack the role of NPD committee recommendations we conduct a mixed methods, three pronged analysis

We look at top down policy perspectives, bottom-up community understanding and business perspective

POLICY PERSPECTIVE – TOP DOWN

- **Examine the ongoing conversations on data sharing legislations in India** such as personal data protection, agri-stack.
- **Deep dive into the mandatory data sharing framework** as recommended by NPD Governance Framework, specifically, its impact on agriculture.

COMMUNITY LEVEL PERSPECTIVE – BOTTOM UP

- **Understanding Digital Green's work** as a consent manager, the digital architecture adopted by them particularly digital wallets, to ensure agency of communities over their data.
- **Understanding Jaljeevika's role as a data steward; methodology used by them to** encourage the voluntary participating of the farming community to share their data by gaining their trust.

BUSINESS/PRIVATE SECTOR PERSPECTIVE

- **Analysis of the possible impact of the NDPR** on existing agricultural businesses and start-ups – how it will affect innovation, intellectual property rights, compliance costs and competition law.
- **Consequences on the fundamental right to privacy** particularly now that businesses may have to mandatorily share data of collected from communities.

We also interviewed researchers, academics and small businesses in the agricultural sector to further hone our research

Figure 2: Analysis approach

04

Policy Analysis

4.1.

Relevant data sharing policies

Over the years, Indian policies have prioritised the setting up of open data platforms, allowing public access to data, and making the data collection, analysis and sharing process transparent. Set out below is an overview of relevant policies that address data sharing:

National e-Governance Plan (2006)[28]

This was an initiative by the Indian Government to make all government services available digitally. Taking cognisance of the e-governance plans implemented across the country, this policy positions itself as an integrated digital platform that provides infrastructure to connect stakeholders across the country, including those from the most remote villages. The end goal of this top-down initiative is to bring public services closer to citizens.

National Data Sharing and Accessibility Policy (2012)[29]

Approved in 2012, this Policy was enacted with the objective of easing public access to non-sensitive data generated using public funds by various branches of both the Central and State governments. The Policy sought to make all such data available in both human and machine-readable format. In pursuance of this objective, the open-government data platform India was launched[30]. The policy also contemplates a pricing mechanism for data decided by data owners and the Government.

Personal Data Protection Bill (2018)[31]

The PDP Bill is the most critical legislation in India with regard to personal data sharing. It seeks to regulate the collection, analysis, and sharing of personal data, both sensitive and non-sensitive, of individuals. It accords rights to data subjects in this regard (such as the ability to confirm or withdraw consent to use data), the grounds on which data of an individual can be processed, and also provides for the setting up of an Indian Data Protection Authority. While the first draft of the legislation was presented in 2018, a revised draft was tabled in Parliament in 2019 and made open for public comments. The Bill is yet to be approved by Parliament.

Image source: https://www.pexels.com/photo/farmland-1287475/
[28] Available at https://www.meity.gov.in/divisions/national-e-governance-plan
[29] Available at https://dst.gov.in/national-data-sharing-and-accessibility-policy-0
[30] Available at https://data.gov.in/government-open-data-license-india
[31] Available at http://164.100.47.4/BillsTexts/LSBillTexts/Asintroduced/373_2019_LS_Eng.pdf

National Data and Analytics Platform (2020)[32]

Proposed by India's nodal policy think tank, NITI Aayog, the platform is intended to democratize access to Government data to make it conducive for public consumption. By hosting the latest datasets from various government websites, and presenting them in a coherent manner, the platform will act as a one stop source for users to utilize, analyse and visualize all published Government datasets. The platform will provide tools to users for analytics and visualization[33]. The first version of the platform is expected to be launched in early 2022[34].

Non-Personal Data Governance Report (2021)[35]

This report provides a framework to regulate the collection and sharing of non-personal data, and enable its access across stakeholders. It strives to achieve a balance between protecting the interests of communities from whom data is collected, and promoting the overall development of the Indian economy. This will be further discussed in later sections.

Aside from the policies mentioned above, there are three policy frameworks, specific to the agriculture sector, which are of interest for this report:

Agristack Platform (2020)

The platform is a national/ centralised database containing personal, sensitive personal and non-personal data of approximately five crore farmers from across the country. It will be used for making data-driven, evidence-based decisions in India using information such as self-reported data on cropping patterns, data on encumbrances, land ownership, type of land held, production history and the financial details of the farmer[36].

India Digital Ecosystem of Agriculture (IDEA) (2020)

This will provide value-added digital services for the implementation of digital schemes in the agricultural sector. Small farm holders will be provided better infrastructure and quality agricultural advisories which can help increase their income and improve overall efficiency and governance.

Unified Farmer Service Platform (UFSP) (2021)

This is a central agency that will provide for seamless interoperability of data between public and private systems in the agricultural sector. This will help provide comprehensive delivery of services to the farmer through data exchanges from government schemes[37].

[32] Available at https://analyticsindiamag.com/what-is-indias-national-data-and-analytics-platform-ndap/
[33] "NITI Aayog Releases Its Vision for the National Data and Analytics Platform", Press Information Bureau (January 23, 2020) retrieved on September 2, 2021 from https://pib.gov.in/PressReleaseIframePage.aspx?PRID=1600370
[34] "Niti Aayog holds meet with 60 think tanks", Financial Express (August 11, 2021) retrieved on September 2, 2021 from https://www.financialexpress.com/industry/niti-aayog-holds-meet-with-60-think-tanks/2308240/
[35] Available at https://ourgovdotin.files.wordpress.com/2020/07/kris-gopalakrishnan-committee-report-on-non-personal-data-governance-framework.pdf
[36] "The AgriStack: A Primer", Internet Freedom Foundation, retrieved on August 2, 2021 from https://internetfreedom.in/the-agristack-a-primer/
[37] National e-Governance Plan in Agriculture (NeGPA): Towards the Mission of Digital Agriculture" retrieved on August 2, 2021 from https://pib.gov.in/PressReleasePage.aspx?PRID=1697526

4.2.

General critiques of India's data sharing policies

While the policies are all based on the same underlying principle – making data accessible for public good – they are disparate, and do not take into account institutional capacities required to help them succeed. There have also been numerous criticisms surrounding the approach adopted by the committees responsible for drafting these policies. For instance, the drafting process across the board was exclusionary as opinions of primary stakeholders were either not accounted for or were restricted to a select few. Calls for comments put out for the PDP Bill[38] and NPDR were limited to individuals and organisations with internet access and the ability to understand the language used in the provisions. This top-down method of policymaking fails to consider the nuances of different sectors, which leads to stakeholders harbouring feelings of mistrust regarding the system.

The process often determines some of the more structural aspects and content of the policy. Most of the policies mentioned above presume the existence of salient features like digital literacy, information symmetry, and internet infrastructure, which is not reflected on the ground (as has been corroborated by expert interviews). As a result of this presumption, the policies do not have provisions for making sufficient investments to develop these competencies. This in turn acts as a contributing factor for the slow growth of India's digital economy despite its potential[39]. Without the adequate development of these competencies, it is impossible to lay the groundwork needed to execute the policies, and broader, more inclusive consultations could help highlight these issues.

The approach of the Telecom Regulatory Authority of India (TRAI) in drafting India's net neutrality policy could serve as a guiding light in this regard. The TRAI accounted for the interests of all stakeholders and followed best practices in the process to draft India's net neutrality policy and this was reflected in the quality of the policy, which is regarded as one of the most progressive net neutrality policies in the world[40].

The PDP Bill's state of limbo in the Parliament points to the most problematic situation in India with regard to data sharing – India does not have a legislation to regulate personal data sharing and protection. Despite this, data sharing frameworks are being developed and implemented, and the risks posed to individual personal and sensitive information are immense. As a result, there is also no way to hold government and businesses accountable for their actions and mismanagement of data. The lack of a personal data protection regulation along with poor processes and disconnected content makes the data sharing landscape in India poorly thought out and haphazardly implemented.

Image source: https://www.pexels.com/photo/farmland-1287475/

[38] Feedback on Draft Personal Data Protection Bill - https://www.meity.gov.in/content/feedback-draft-personal-data-protection-bill

[39] "Digital India has the potential to create a $1 trillion digital economy", Dhotre S., The Energy and Resources Institute (November 7, 2019) retrieved on July 30, 2021 from https://www.teriin.org/article/digital-india-has-potential-create-1-trillion-digital-economy

[40] "Timeline: How the net neutrality debate opened up in India", Business Today (December 2014) retrieved on August 5, 2021 from https://www.businesstoday.in/latest/economy-politics/story/timeline-how-the-net-neutrality-debate-opened-up-in-india-88956-2017-11-28

4.3.

Agriculture sector-specific critiques of India's data sharing policies

There are also significant points of concern with the policies when viewed from an agriculture sector-specific angle. The policies do not ensure last mile access to technology and the ability of people to generate data and use it in meaningful ways. Even though information is collected from the agricultural sector, specifically farmers, none of the policies specifically talk about insights being made accessible back to the farming community. The data is made available only online, which is highly problematic given the poor internet penetration (especially in rural areas)[41] and extremely low levels of digital literacy among farmers[42].

Beyond questions of access, there are no provisions for confirming the consent of farmers while sharing and re-sharing data on large public platforms[43]. Farmers are thus unaware of how their data is being used or whom it is being shared with.

The policies also do not address the existing issues of poor quality of datasets and how they plan to resolve this before sharing the data for public purposes. The issue with the credibility and quality of the datasets is well known, but the policies do not account for this in any manner, thereby leading to unreliable datasets being shared for analysis.

The pricing policy provisions of the policies also do not specifically address compensation for the farmer community, which is a crucial incentive for farmers to share data. Additionally, there is no provision for the protection of IP rights of farmers while sharing data, meaning that the farmers are unable to derive any economic benefits from sharing their data. Lack of a personal data protection law can render individual and community data susceptible to infringement.

These problems are emblematic of another major concern with the top-down nature of policy making. Without adequate consultation from relevant stakeholders, especially farmers, the potential risks and harms to farmers are not addressed.

Finally, concerns of the agricultural sector are compounded by the fact that India still does not have data protection legislation. Data-driven innovation cannot be scaled without adequate privacy safeguards and gaining users' trust. Strict regulations such as localisation mandates neither ensure privacy nor support innovation and entrepreneurship. Without a law on privacy, data subjects can only approach the courts in case their data is hacked or rights are infringed, which can be financially burdensome or procedurally difficult. Businesses cannot be held legally accountable as they are still not mandated to set up data protection authorities to address personal data grievances of users.

Image source: https://www.pexels.com/photo/farmland-1287475/

[41] "Connectivity gets better but parts of India still logged out", Kawoosa V. M., The Hindustan Times (August 14, 2020) retrieved on September 3, 2021 from https://www.hindustantimes.com/india-news/connectivity-gets-better-but-parts-of-india-still-logged-out/story-VSqXriMdGUudWb7eBcWzjN.html

[42] "The digital dream: Upskilling India for the future", Mothkoor V. & Mumtaz F., Ideas for India (March 23, 2021) retrieved on September 3, 2021 from https://www.ideasforindia.in/topics/governance/the-digital-dream-upskilling-india-for-the-future.html

[43] "#PrivacyOfThePeople - Agristack and Farmers' Issues", Internet Freedom Foundation (June 30, 2021) retrieved on September 3, 2021 from https://internetfreedom.in/privacy-of-the-people-agristack-and-farmer-issues/

According to experts, another potential issue that has to be addressed is the technology lock-ins by businesses that force farmers to utilise the machinery of businesses and services related to it, irrespective of their consent. These all-or-nothing contracts leave them with little say in the matter.

The government is collating the personal and sensitive personal data of the farmer community in the absence of privacy legislation without confirming consent, in ways that may act against their interests. While this is a clear violation of their right to privacy, it is difficult to contest as it is being done in the name of promoting public good.

The Agristack platform (see Figure 3) is an aggregated example that is reflective of all these problems. According to this explainer[44], the platform is a national/centralised database containing personal, sensitive fiscal and non-personal data of approximately five crore farmers from across the country. It will be used for making data-driven, evidence-based decisions in India using information such as self-reported data on cropping patterns, data on encumbrances, land ownership, type of land held, production history and the financial details of the farmer. While experts we spoke to agreed that there is value in creating a centralised database as it helps to realise disbursements, the system must ensure that farmers retain control of their data. They pointed out that these datasets are looked at from a market-oriented perspective, as in the IDEA[45] paper, rather than an individualistic perspective.

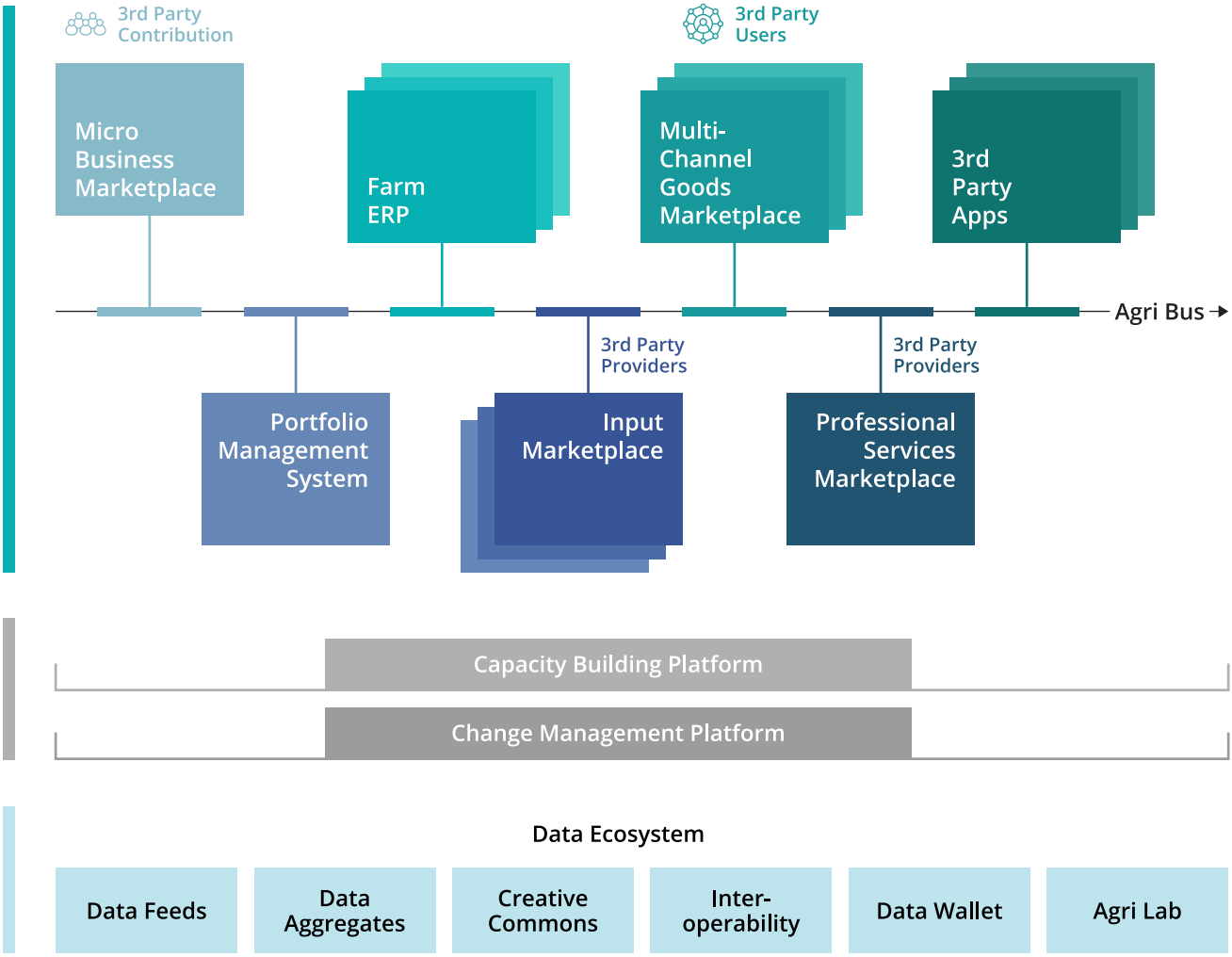


Figure 3: Overview of Agristack

[44] "The AgriStack: A Primer", Internet Freedom Foundation, retrieved on August 2, 2021 from https://internetfreedom.in/the-agristack-a-primer/
[45] "Consultation Paper on IDEA" retrieved on September 2, 2021 from https://agricoop.nic.in/sites/default/files/IDEA%20Concept%20Paper_mod01062021_1_0.pdf

In April 2021, 55 organisations drafted a letter to the government explaining their concerns over the platform and providing suggestions to deal with the issues[46]:

Lack of consultation

Primary stakeholders – India's farmers – were not consulted for the design of Agristack. There is no evidence of their interests being included in the policy, making the adoption of techno-solutionism an unsustainable option.

Questions of harm unaddressed

Unintended harms and exclusion of farmers from services such as financial services are not adequately addressed by the policy. Additionally, the overdependence on artificial intelligence to make sensitive, life-impacting decisions without any transparency about the process puts the farmer at risk.

Questionable data quality

There is no information to attest to the fact that the information in the database is representative of all farmers. As pointed out by one of our interviewees, tenant farmers, for instance, are not included as 'landowners' – this is an issue that is still in conflict. Without clarity on this, they may be denied benefits that they would have gotten had their information been in the database. Crucial decisions are also made with these datasets sans any transparency, which can be extremely problematic.

No clarity on implementation status

After an initial concept note, there has been no follow-up with respect to the status of implementation or review of the policy based on feedback and criticism provided.

Image source: https://www.freepik.com/premium-photo/indian-farm-workers-harvesting-green-coriander-holding-bunch-hands-organic-farm_13506584.htm#page=6&query=crops%20farm&position=11

[46] "ASHA Letter to GoI on Direction & Partnerships of its Digital Push by Ministry of Agriculture", Kisan Swaraj, retrieved on August 2, 2021 from http://www.kisanswaraj.in/2021/05/05/asha-letter-to-goi-on-direction-partnerships-of-its-digital-push-by-ministry-of-agriculture/

Across policies and regulations, India has not prioritised a few key metrics that can build a robust data sharing ecosystem. These issues are both procedural and structural and need to be addressed to ensure that data sharing in agriculture creates positive impact for all stakeholders. If the core issues are left uncontemplated, it might hamper the overall objective of these policies – easing access to data by the public, while creating the best value out of it.

REGULATION / POLICY	OBJECTIVE	CONCERNS
National e-Governance Plan (2006)	• Reliable access to internet, large scale digitization of land records • Build a massive, country wide technology infrastructure reaching the most remote village	• Presumes existence of digital literacy in the country • Top down framework - policy non consultatively drafted
National Data Sharing and Accessibility Policy (2012)	• Shareable, non sensitive data collected using public funds to be made available to the public	• Government non-consultatively prepared 'negative list' - data that cannot be shared with the public • Data owners asked to come up with a 'pricing policy' for the data but "data owners" and profit undefined
Personal Data Protection (PDP) Bill (2018)	• Protection of the personal data of individuals • Establishes roles and responsibilities for the data collection and process	• Does not examine the possibilities of stakeholders who may refuse to give their consent to share personal data • Consent is not structured as dynamic • Focussed on protection, not empowerment
National Data Analytics Platform (2020)	• Visualises government published datasets (Central and State) in an accessible, user friendly format	• Target audience for platform is not defined as farmer groups • Consultation with primary group - farmers - was not undertaken
Non Personal Data (NPD) Governance Framework (2020)	• Defines what constitutes non personal data • Establishes new regulatory bodies to oversee the NPD collection and processing mechanism	• Makes data sharing mandatory for public good, but lacks nuance • The closest possible representation of the community is the data trustee - this however has been restricted to being either a not for profit organisation or government body
Agristack Platform (2020)	• National / centralised database containing personal, sensitive personal and non personal data of approx. 5 crore farmers across the country • Enabled data-driven, evidence-based decisions using information on self-reporting data	• Questions of unintended harms and exclusion of farmers from services not addressed • Questionable data quality - no information to show that database is representative of all farmers
InDEA Digital Ecosystem of Agriculture (IDEA) (2020)	• Provision of value-added services for implementation of digital schemes in the agri sector • Provide access to high-quality data in agri sector to stakeholders	• Lack of consultation with farmers in drafting policy • Existing problems unaddressed, including protection of privacy of farmers
Unified Farmer Service Platform (UFSP) (2021)	• Central agency to provide interoperability of data between public and private systems in agriculture sector	• Policy framed in a top-down manner non consultatively • Overdependance on technology, lack of human interface

Figure 4: Overview of data sharing policies & regulations (proposed and enacted)

05

Non Personal Data Framework

5.1.

Overview of the NPDR

The Report by the Committee of Experts on Non-Personal Data Governance Framework (NPDR) serves as the first report in India setting out a framework for the governance of non-personal data. As mentioned previously, the report is the first policy document in India that prioritises the interests of communities, the originators of the raw NPD.

In deliberating personal data protection legislation in India, the members of the Srikrishna Committee (the expert committee set up to draft a personal data protection legislation) also noted the importance of community data – aggregated information that was used for understanding consumer behaviour, public preferences, and making decisions to promote social welfare. This questioned the idea of individual control over privacy, considering the multitude of data subjects who contributed to this database.

In response, the Government set up a committee to legislate on the notion of 'collective privacy' in order to accord protection to the particular community from whom the data stemmed. Led by Kris Gopalakrishnan, co-founder of Infosys, the Committee of Experts on Non-Personal Data Governance Framework decided to focus on the regulation of non-personal data, which it recognised to be a national economic asset[47]. This Committee's report recognises the different elements of non-personal data – derived data, e-commerce data, anonymised data, etc. – all of which are critical for drafting policies and delivering public services to citizens[48].

The Committee had released the first version of its report in July 2020[49], and invited comments from stakeholders. The Committee then released a revised report in December 2020, taking into account the public comments[50]. The NPDR recognises the increasing importance of non-personal data, particularly its contribution to economic, social and public value. NPD, in particular, has taken 'centre stage' in the working of core technological businesses, and has helped address various public administration issues[51].

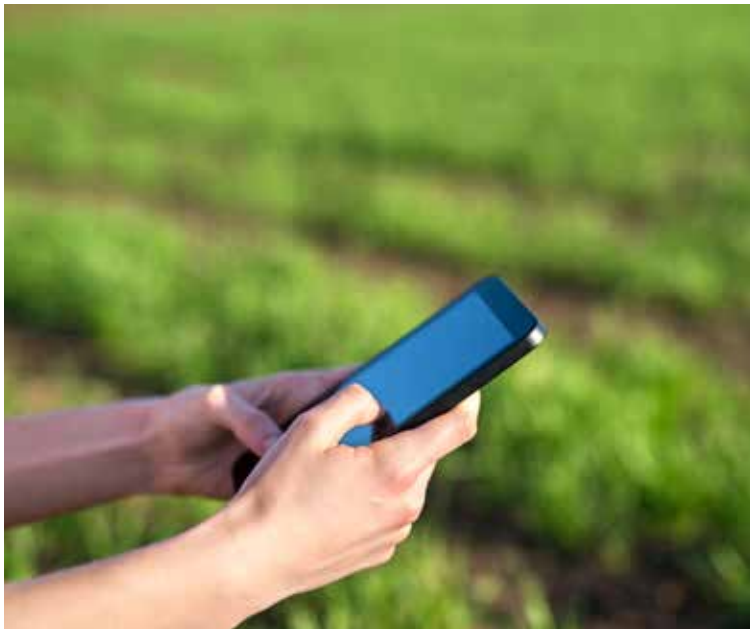


Image source (top): https://www.pexels.com/photo/photo-of-corn-field-3066814/
Image source (bottom right): https://www.freepik.com/free-photo/growing-crops-with-modern-technologies_11133950.htm#page=3&query=crops%20farm&position=44
[47] "Indian Govt Forms Committee To Recommend Governance Norms For Non-Personal Data, Infosys' Gopalakrishnan To Head It", Agrawal A., Medianama (September 16, 2019) retrieved on August 10, 2021 from https://www.medianama.com/2019/09/223-meity-non-personal-data-committee/
[48] Official notification announcing the Constitution of a Committee of Experts to deliberate on the data governance framework: https://www.meity.gov.in/writereaddata/files/constitution_of_committee_of_experts_to_deliberate_on_data_governance_framework.pdf
[49] "India: Meity releases revised report on non-personal data framework for public consultation", Data Guidance (January 4, 2021) retrieved on August 10, 2021 from https://www.dataguidance.com/news/india-meity-releases-revised-report-non-personal-data
[50] The Committee classifies non-personal data as any information that does not fall within the ambit of 'personal data' (as defined under the Personal Data Protection Bill) or data that is without any personally identifiable information
[51] Page 5, NPDR report

The following are some of the key provisions from the Committee's report:

5.1.1. Definition and classification of NPD

The Committee defines NPD as any information that is not personal data (as defined in the PDP Bill) or data that is without any personally identifiable information. The data cannot be used to identify a natural person, for example, data about the weather, or information collected from sensors on industrial machines. It may also include data that was initially personal data, but which was later aggregated or anonymised to the extent that individual identifiers are no longer associated with it. NPD is further classified into three categories:

Public data

Data generated by the government or through publicly funded works such as datasets or documents published by ministries or their departments.

Community data

Raw data sourced from groups of individuals "who are bound by common interests and purposes and involved in social and/or economic interactions".

Private data

Data generated or collected by businesses or private players in the market.

The focal point of the NPDR is community data, with the Committee trying to find the best ways to derive value from it, while protecting the interests of the data principals (the individual, community or company from whom NPD is collected).

Image source: https://www.freepik.com/premium-photo/indian-goat-street-rural-india_12189910.htm

5.1.2. Formation of institutions for regulating NPD sharing

In order to account for value creation and community protection, the report creates new institutional mechanisms that will be formed within the existing market system.

Data custodians

which includes government or private organisations that collect, store and process NPD on behalf of the data principal.

Data trustees

act as representatives for community NPD, and can either be a government body, private organisation or even a coalition of the community members themselves.

Data businesses

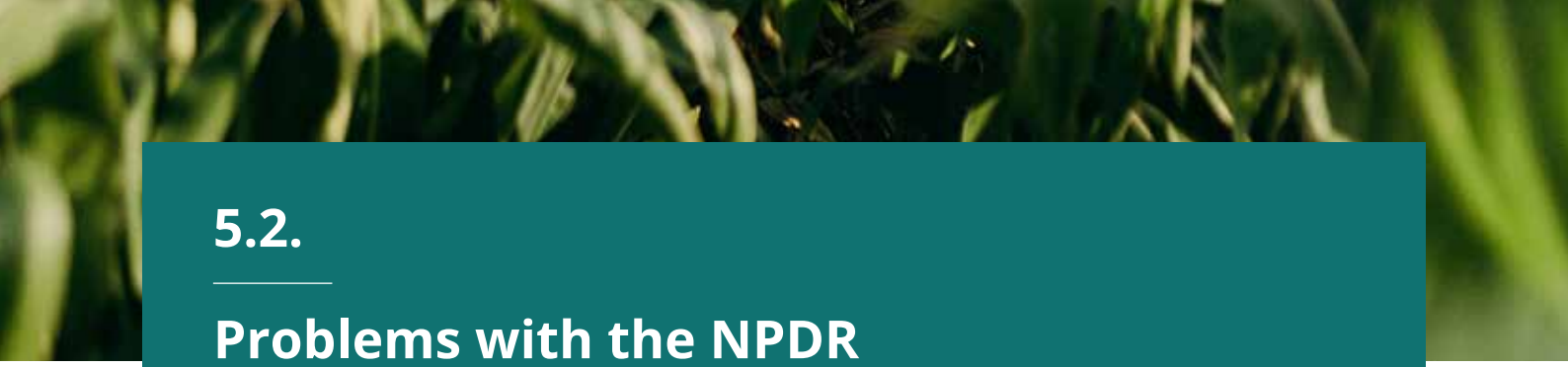
government or private organisations that collect, process or store NPD beyond the threshold that will be specified by a regulator. Data businesses are classified depending on their 'data drivenness' (a term that has yet to be defined) in the market. The report also claims that the demand and supply side will be factored into this decision, though it does not explain how. This is a part of the Committee's plan to challenge the existing monopolistic market system in India – big businesses will have to share their data in return for a nominal amount.

Image source: https://www.freepik.com/premium-photo/indian-goat-street-rural-india_12189910.htm

5.1.3. Optimising value creation from high value datasets (HVD)

The sharing of data has been made mandatory for a new category of data called HVD. According to the report, this includes information that is beneficial to the community and is needed for promoting public good. HVD are further classified into three categories – raw data (for example, census data of a citizen), aggregated data (for example, agglomerated orders of all consumers from an e-commerce website), and inferred data (information derived after analysing datasets using company algorithms). A data requester can seek access to this HVD from the data trustee, who will be registered in India and is responsible for maintaining the data. Only organisations, and not individuals, will be granted access to HVD, and will have to pay nominal charges to the data trustee for this. The HVD will have the same protections and storage requirements as personal data, considering they are prone to harms such as de-anonymisation, which can affect the data principal. For this purpose, adequate digital architecture will have to be set up, and cloud servers constantly tested for compliance. An advisory body consisting of representation from academia and industry will be set up to overlook the implementation of these protections, and to make recommendations for technological interoperability, governance standards, sharing frameworks, etc.

Image source: https://www.freepik.com/premium-photo/unidentified-indian-farm-worker-harvesting-green-coriander-holding-bunch-hands-organic-farm_12081129.htm#query=crops%20farm&position=2#position=2



5.2.

Problems with the NPDR

While the intent of the NPDR is laudable, there remain a number of issues with the content of the policy and the process adopted in its promulgation. To begin with, the NPDR seems to be predicated on a perceived lack of public good initiative from all private actors, with little mention of G2B sharing. This approach only serves to exacerbate the existing trust barrier between the public and private sector. Moreover, the requirement for mandatory data sharing is likely to create both regulatory and compliance burdens, thereby increasing compliance costs for businesses (big or small). While the NPDR has a focus on preventing accumulation of power by big businesses and preventing monopolies, the compliance costs that will be imposed can hamper start-ups from entering the market, and existing firms, especially small companies, from growing.

The NPDR also defines a number of concepts which are broad and require clarity, the definitions of NPD and communities being two prime examples. This room for interpretation can lead to misuse of powers by authorities. Definitions, like that of 'data trustees', also create grounds for conflicting interests, especially in the sphere of mandatory data sharing with the government, as government bodies qualify as data trustees themselves[52].

In addition to the broad definition of a 'community', there is a failure to address critical community related issues such as the exercise of community agency over data and confirming community level consent. There are also other issues around consent, and how community interests may balance with public interests and how the NPD and PDP will interface with one another. Community-based organisations and civil society organisations will have an additional responsibility of navigating complex data sharing systems on behalf of communities and ensuring that use of NPD does not harm communities. Further, while the role of data trustees has been identified, communities may not have adequate space to negotiate on aspects of mandatory data sharing which will infringe upon data rights. Businesses invest large amounts to collect and analyse data, with the aggregate data they collect and analytics from the data belonging to them (with the exception of information necessary in the case of public emergencies). They also own and control IP over algorithms used for processing. The data sharing system under the NPDR mandates the sharing of this data without any compensation. Businesses will thus be forced to share proprietary data at a loss. Additionally, mandatory frameworks in data related industries across jurisdictions have seen poor levels of compliance. The Australian Competition and Consumer Commission's proposals, which would, inter alia, mandate big tech companies like Google to share users' anonymised personal data with advertisers, were met with backlash, with Google threatening to stop provision of services in Australia[53]. Similarly, in India, companies like Whatsapp and Google have approached Courts over new Information Technology Rules which mandate social media intermediaries to share

Image source: https://www.pexels.com/photo/photo-of-cornfield-3066814/

[52] "India's Latest Privacy Guidelines Still Need Fixing", Mahesh S., Freedom Gazette (January 24, 2021) retrieved on August 10, 2021 from https://www.freedomgazette.in/2021/01/indias-latest-privacy-guidelines-still-need-fixing/

[53] "Australia take on Google advertising dominance in latest Big Tech fight", Kaye B., Reuters (January 28, 2021) retrieved on September 2, 2021 from https://www.reuters.com/article/us-australia-media-regulator-idUSKBN29X02X

information about 'originators' of 'unlawful' messages[54]. There has already been a push back from big tech firms such as Amazon and Facebook against mandatory sharing of NPD[55].

The NPDR is an umbrella policy that cuts across all sectors and pushes for the introduction of a mandatory data sharing framework, for which it has faced criticism from experts. Different sectors and different data types each have their own set of considerations which require nuanced approaches to sharing and management. The NPDR's mandate leaves little room for this. However, the Government is also parallelly working on sectoral governance of non-personal data, in policies like the National Digital Health Mission (NDHM). Such parallel processes, done without synergy between the concerned drafting groups, or the policies themselves, serve to exacerbate the issue of multiple, conflicting policies on data sharing governance. There is a need for alignment between the NPDR and individual sectoral policies on NPD[56].

REGULATION / POLICY	CONSULTATIVE *Developed in consultation with focus community*	BOTTOM-UP *Regulation discussion led by sector/community*	INCENTIVE FOCUSED *Monetary & instrumental incentives to participate*	INNOVATION FIRST *Prioritises data-led innovation*
National Data Sharing & Accessibility Policy (2012)	Somewhat — Consultations only with government agencies	No — Government initiated and led	No — No clear incentives to participate in open data	Somewhat — Innovation is claimed but pathways are undefined
National Data Analytics Platform (2020)	Somewhat — No explicit consultations mentioned	No — Government initiated and led	No — No clear incentives to participate in the platform	Somewhat — Innovation is claimed but pathways are undefined
Personal Data Protection (PDP) Bill (2018)	Somewhat — Consultations only through comments on draft	Somewhat — Civil society articulated need for data protection; government led	Somewhat — Compliance only incentives	Somewhat — Protection framework only
Non Personal Data (NPD) Governance Framework (2020)	Somewhat — Consultations only through comments on draft	No — Government initiated and led	Somewhat — Public value is seen as an incentive but undefined	No — Mandatory sharing will stifle innovation
Agristack	Somewhat — No consultations called for during development and pilot	No — Government initiated and led	Somewhat — Benefits to farmers highlighted but undefined	Somewhat — Innovation is claimed but pathways are undefined

Yes Somewhat No

Figure 5: Process review of data sharing regulation and policy

Image source (top): https://www.pexels.com/photo/photo-of-cornfield-3066814/

[54] "Explained: WhatsApp's arguments to fight traceability clause in IT Rules 2021", Dhapola S., The New Indian Express (June 2, 2021) retrieved on September 2, 2021 from https://indianexpress.com/article/explained/whatsapp-india-it-rules-traceability-clause-case-explained-7331039/

[55] "Exclusive: India data-curb plan 'anathema', U.S. tech giants plan pushback", Kalra A., Reuters (August 9, 2020) retrieved on September 2, 2020 from https://www.reuters.com/article/us-india-data-exclusive/exclusive-india-data-curb-plan-anathema-u-s-tech-giants-plan-pushback-idUSKCN2550KA

[56] "India needs a digital health mission. But it also needs data privacy law to ensure it works", Patnaik I. et. al., The Print (August 21, 2020) retrieved on August 19, 2021 from https://theprint.in/ilanomics/india-needs-a-digital-health-mission-but-it-also-needs-data-privacy-law-to-ensure-it-works/486111/

5.3.

Agriculture sector-specific problems of the NPDR

We analysed the possible implications the NPDR would have on the agricultural sector, by looking at its impact on two key stakeholder groups – farmer communities and small businesses working in agriculture. Interviews with grass-roots level organisations, and consequent analysis, gave us insights into the adverse implications that the NPDR can have on the agricultural sector. The top-down policy drafting mechanism adopted in framing the NPDR has led to the exclusion of primary stakeholders from the discussion, and as a result the NPDR fails to take into account their concerns and considerations, particularly the elements required to ensure their participation in the system.

Image source (top): https://www.pexels.com/photo/photo-of-cornfield-3066814/
Image source (bottom): https://www.freepik.com/premium-photo/indian-farmer-working-traditional-way-with-bull-his-farm-indian-farming-scene_9170693.htm

5.3.1. NPDR concerns for farmer communities

Given the low levels of digital literacy of farmers in India, significant investments in improving digital literacy are needed to ensure that the benefits of digitization schemes, including under the NPDR, reach farmers. This low level of digital literacy sows mistrust in farmers towards the framework, as they see no sufficient benefit accruing to them.

The mandatory data sharing under the NPDR can also have a significant impact on the IP rights of farmer communities. While open access to data is crucial for development of public good, it must not come at the cost of the IP rights of farmer communities. The existing IP regime in India is inadequate to help farmers protect their IP rights. The Copyright Act, 1957 does not accord copyright for raw data. As a result, there is no protection for farmers who collect their own data about their farms. Further, copyright is threatened when data is made publicly available as the farmers then do not have any remedy owing to exceptions within the act which take precedence, such as that of fair use. India does not have legislation on trade secrets, and protection for trade secrets are granted by Courts on a case by case basis. Farmers are not updated on jurisprudence regarding trade secrets and confidentiality, and approaching courts for protection can be a time and resource intensive process. Finally, while the Protection of Plant Varieties and Farmers Act, 2001 accords protection to new plant varieties and genetic material of traditional communities and mandates benefit sharing and compensation to farmers by breeders for commercial use, it has seen a difficulty in compliance and farmers face obstacles in the practical application of this legislation. The NPDR, in mandating data sharing, opens up further avenues for the IP of farmers to be used without any compensation due to them. The NPDR framework does not address issues of benefit sharing or compensation either to farmers or to businesses.

The definition of a 'community' under the NPDR is both broad and vague, leading to confusion on who constitutes a valid community under the NPDR. The NPDR adopts a homogenised conception of communities as monolithic entities bound by the same interests. However, the reality is very different and farmer communities are uncertain as to how varying interests will be reconciled. The NPDR has sought to address this in part through the role of a Data Trustee. However, the opaque, ambiguous characterisation of a Trustee, a lack of transparency of their powers and responsibilities engenders a relationship of subordination between the community and the Trustee, opening the former to exploitation by the latter. This leads to further confusion at a grass roots level on how farmer communities will be able to protect their NPD.

Additionally, the NPDR does not clearly address the difference between community non personal data, and public and private non personal data. While community data cuts across personal data and NPD, the NPDR typifies it as a third type. This leads to further confusion on what constitutes NPD and the boundaries between consent for private data and consent for community NPD. This framework will only further compromise the already fragile privacy rights of farmers.

Image source: https://www.freepik.com/free-photo/farmers-harvest-barley-happily_4284017.htm#page=1&query=farmers&position=28

5.3.2. NPDR concerns for businesses in the agriculture sector

About 80% of the agriculture sector is driven by the private sector[57]. Small businesses create significant value for farmer communities, making it crucial to continue this robust engagement with them and the private sector already utilises data, making large contributions to public value and economic growth. The NPDR does not recognise the innovation brought by these businesses to the sector. The provisions of the NPDR might disincentivise them from participating in the framework, thereby reducing their contribution to the economy.

To begin with, there is a lack of clarity about the threshold for data driven businesses to mandatorily share data. The NPDR distinguishes between businesses based on their 'data drivenness', but it does not define what this specifically constitutes. As a result, small and medium businesses are unsure of whether they fall into this category. This becomes particularly problematic given that the NPDR imposes a number of compliance requirements. Businesses, particularly small agricultural businesses and start-ups, may not be aware of the nuances of the NPDR and the compliance requirements therein. In addition to this, many businesses, and start-ups, existing consent managers and data trustees/stewards (civil society organizations) may not be able to withstand the increase in business costs needed to comply with the NPDR, resulting in a high barrier to entry. The NPDR also leaves the prices for data processing for businesses to negotiate, which can be a problem for small businesses and start-ups, because of existing power imbalances. It is therefore imperative that the Government allocate the time and funds needed to educate them, and invest in resources to support them in compliance.

Businesses invest large amounts to collect and analyse data. The NPDR says that businesses will only be compensated for data processing, as data collection is a part of their regular course of business. Such mandatory sharing of data without compensation serves to disincentivise businesses from participating.

The NPDR also requires businesses to share inferred data and metadata, opening an avenue for current methods of working and future plans to be revealed. This can have negative impacts on sectoral competition, and hamper innovation among businesses.

Overall, the NPDR aims to ease the process of sharing and unlock value for stakeholders, especially the community that generates data. While the regulations are a starting point for discussions on non personal data sharing, there are still crucial gaps that need to be addressed for it to work effectively. There are several new concepts in the report that need to be expanded, nuanced and tested, roles of different stakeholders need to be clarified, and aspects such as mandatory sharing need to be revisited, especially with regard to their impact on the data sharing ecosystem. Beyond this, the application of the policy from a sectoral lens is critical to understand how things might unfold on the ground.

Image source: https://www.freepik.com/premium-photo/indian-farmer-golden-wheat-field_8179414.htm#page=3&query=farmers&position=10

[57] "How private sector is helping cultivators with technology, buyback and improving their social standards", Sally M., The Economic Times (December 8, 2017) retrieved on September 3, 2021 from https://economictimes.indiatimes.com/news/economy/agriculture/how-private-sector-is-helping-cultivators-with-technology-buyback-and-improving-their-social-standards/articleshow/61989717.cms

Even sector specific agricultural businesses focused on farmer value are reluctant about participation in the NPD regime

ISSUES	EXPLANATION
DIGITAL LITERACY OF FARMERS	• Only 7-8% of total farmers are digitally literate; Government needs to invest in this aspect for them to **benefit from digitization schemes.** The ecosystem needs to be developed to focus on data transmission and data consumption.
MANDATORY DATA SHARING SYSTEM	• Businesses invest in large amount to collect and analyse data – aggregates and analytics belongs to them (exception: public emergencies); Owns and controls IP over algorithms used for processing – **sharing this data with loss of profits unfair.** • Only data generated with public money to be made open
LACK OF INCENTIVES TO ENCOURAGE DATA SHARING	• For businesses - Government pays for the data (payment = cost + income for data creator) – **compensation needed** • For farmers – Adherence by businesses and Government to data privacy laws **(ex. While taking consent, providing complete information)**
LIMITED PROTECTION OF BUSINESSES' IP RIGHTS	• Possibility of IP rights infringement – particularly trade secrets of businesses because **it is not regulated** • Misuse of farmers data by competitors – loss of business advantage, price manipulation – **anti trust effects to be penalised**

Small businesses are creating significant value for famrer communities and there is a need to continue a robust engagement with them – the NPD does not recognise the innovation brought by SMBs to the sector

Figure 6: Concerns of Agricultural Businesses with the NPDR

5.4.

Role of Ministry of Agriculture and Farmers' Welfare

The Ministry of Agriculture and Farmers' Welfare is the central point for sectoral data collection activities. The Ministry collates information from regional, state and national levels, including weather data and soil data, data on 'structural aspects of operational landholdings' via the agricultural census conducted every five years, data on success rates of government agricultural schemes like PM Kisan and Soil Health Card, and data on consumption of various agricultural inputs according to major size-groupings of operational holdings collected through its input surveys.

A detailed reading of the reports from the Ministry highlights a number of flaws in the Ministry's approach to data collection. The Ministry does not explain the methodology adopted to collect the raw data and the means of analysis, representing the final outcome in its report. Ad hoc methods appear to be used for varying categories of populations and the quality of data is considered questionable. This was also confirmed in the interviews with small agricultural businesses who mentioned a reluctance to depend on this data – citing the lack of transparency around the methodology making the data quality highly questionable[58].

Personal and non-personal data are collected from farmers and presented in government reports in an aggregated and anonymised form. The reports do not make reference to any data protection policy nor is there any information about adherence to privacy protection regulations. Without legislation to protect this information, data subjects will continue to be exposed to risks of infringement on privacy, and will not have a redressal mechanism. With the PDP Bill still under discussion in Parliament, farmers are unable to exercise any rights as data subjects[59]. The consent mechanisms adopted in the data collection and the anonymisation standards followed in anonymising the data are also unclear. When it comes to non-personal data – which has been classified by the NPD Committee as any data that has been anonymised – there are risks posed to the data subjects, such as the harms that may occur due to re-identification, either due to poor anonymisation technology or association of the data with other datasets[60].

While it is mentioned that the information collected is used to evaluate existing agricultural programmes (based on which necessary improvements are made to increase crop production and productivity, and to provide an advisory on issues such as plant protection, residue management and soil fertility)[61] there is little information on the criteria applied in using the data to evaluate existing agricultural programs.

While making these reports available online for public access might fulfil the objective of

Image source: https://www.pexels.com/photo/photo-of-cornfield-3066814/

[58] Excerpts from our interviews carried out with small agricultural businesses based in India. The names of these businesses have been kept confidential upon request.

[59] "JPC proposes to expand ambit of personal data protection bill", Varma G., Mint (November 25, 2020) retrieved on August 19, 2021 from https://www.livemint.com/news/india/jpc-proposes-to-expand-ambit-of-personal-data-protection-bill-11606269336962.html

[60] NPDR, Page 10

[61] "Results-Framework Document (RFD) 2013-14", Ministry of Agriculture and Farmers Welfare, available at: https://agricoop.nic.in/sites/default/files/RFD%20_2013-14_14052013.pdf

the government's open data project, it is of no practical use to the farmers. They cannot access these findings, the details of which stem from their raw data, either due to poor digital literacy or lack of internet connectivity.

Finally, data sharing by the Ministry for "public purposes" takes place in an ad hoc and undefined manner, which is also a major concern in the NPDR. This exchange of data is done with three stakeholders:

Private companies

Example, the Department of Agriculture and Farmers Welfare has entered into MOUs with private sector companies such as Star Agribazaar, CISCO, and ITC[62]. While these MOUs have been signed to help with the implementation of the Agristack project, there is no clarity in any of the MoUs regarding how the data sets are to be kept secure. Further, the Non-Disclosure Agreements (NDA) in the appendices of the MoUs are vague on the subject of whether farmers' data shared under the respective MoUs is considered "confidential information", and thus subject to protection under the NDA[63].

Researchers and policymakers

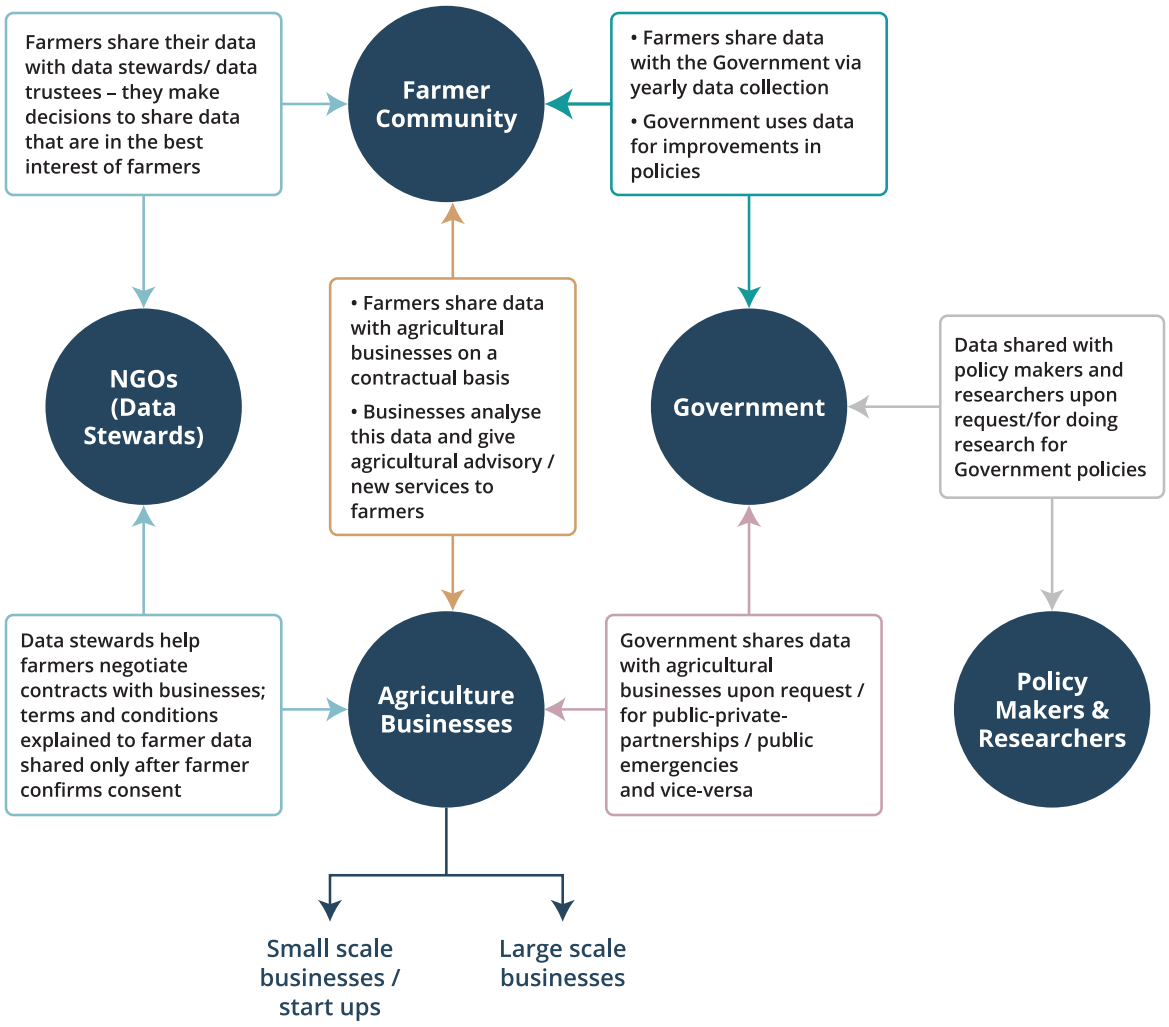
Example, data sharing with the Indian Council of Agricultural Research (ICAR)[64] (an autonomous organisation under the Ministry). This information is then used to draft agricultural policies and legislation. However, researchers have expressed their frustration because of the hurdles they face while trying to access databases[65]. Without the data, they have very little visibility about the overall issues plaguing the sector, which can be a major limitation in their study.

Civil society organisations representing farmer communities

Example, Non-governmental organisations (NGOs) such as Rythunestham. An organic farming advisory based in Hyderabad, it has developed an application in collaboration with the central government that provides individual advisories to farmers[66]. However, there have been issues about the data being outdated, leading to inaccurate information being given to farmers.

Overall, there are several concerns on how the Ministry manages data collection, quality and accessibility as well as the capacity of the government to meaningfully share data for public interest. A lack of data protection policies, opaque data collection methods, complex procedures for accessing data and data sets of questionable quality have resulted in a framework that stifles innovation in the economy while actively disincentivising stakeholders from participation. These issues remain largely unaddressed in new drafts of the data protection and sharing frameworks in India, including the NPDR. While the focus of policies has shifted to the interoperability of data in the agricultural sector, they are pivoted on the objective of economic development. As a consequence of this, farmers' interests have been continuously de-prioritised.

[62] MoU. Ministry of Agriculture and Farmers' Welfare, available at: https://agricoop.nic.in/en/MoU
[63] "MoUs for algorithms and data for profit: 4 new MoUs signed between the Ministry of Agriculture and private corporations, including Jio and Cisco" Internet Freedom Foundation, retrieved on October 4, 2021 from https://internetfreedom.in/algorithms-in-mous-and-data-for-profit-4-new-mous-signed-between-the-ministry-of-agriculture-and-private-corporations-including-jio-and-cisco/
[64] https://icar.org.in/node/173
[65] "There are so many hurdles.' Indian scientists plead with government to unlock COVID-19 data", Pulla P. (May 4, 2021) retrieved on August 2, 2021 from https://www.sciencemag.org/news/2021/05/there-are-so-many-hurdles-indian-scientists-plead-government-unlock-covid-19-data
[66] "Mobile apps are empowering farmers", Extension Digest (December 2017) retrieved on August 2, 2021 from https://www.manage.gov.in/publications/edigest/dec2017.pdf



- **Methodology** for data collection vague; **data protection** policies not specified – **data subjects rights** unknown; no support to protect **IP rights** of farmers
- Analysed data shared on Government website – **inaccessible** to farmer community; **indirect impact** felt on community via new policies/ improvements in existing policies ; farmers **excluded** from policy making

- **Quality of datasets** maintained by Ministry questioned by businesses; **complex procedure** for accessing data discouraging; businesses **excluded** from policy making
- **Limited data** shared with Government, businesses' IP interests given priority; data sharing done at **high costs**, burden on Government

- **Quality of Government** datasets questionable; **complex procedure** for accessing data discouraging for researchers; results of research not shared with farmer community (primary stakeholders)

- Data stewards collect data **haphazardly** – may not be aware of data subjects privacy rights; burden to update consent, understand nuances of business contracts, make ethical decisions in farmers interests
- NGOs share data as stewards – **imbalance of power** – NGOs forced to negotiate; may not have resources to understand complexities of **contracts** with businesses

- **Imbalance of power** between businesses and farmers – unfair provisions in contracts (example – technological lock-ins, excess data sharing, non-purpose specific use of data); **IP rights** of farmers susceptible to infringement
- Agricultural advisory given by businesses to farmers at **high costs**; rights of data subjects given **lesser priority** over business interests

Figure 7: The broken agriculture data sharing ecosystem

OTHER RELATIONSHIPS

NGO/Civil society - Government:

NGOs use Government datasets – nationwide picture of the agricultural sector – quality of datasets questionable, complex procedure for accessing data discouraging

Small scale businesses/start ups – Large scale businesses:

Imbalance of power- start ups, small businesses forced to negotiate for data with big tech; high costs imposed for limited data sharing; might be forced to merge or exit the market

Small scale businesses/start ups – Government:

Start ups/small agri-tech businessess use Government datasets – nationwide picture of the agricultural sector – quality of datasets questionable, complex procedure for accessing data discouraging

Figure 7: The broken agriculture data sharing ecosystem

Image source: https://www.freepik.com/premium-photo/farmers-workers-are-plowing-sowing-agricultural-field-traditional-way-with-help-bulls_9768815.htm#page=11&query=farms&position=25

06

The Need for an Alternative Data Sharing System

6.1.

Challenges of mandatory data sharing frameworks

This system is oriented towards economic development, leaving little scope for promoting innovation and accounting for interests of primary stakeholders. It has been argued that mandatory legislation will be able to ensure compliance by stakeholders through the issuance of penalties. This will ensure access to data by small businesses and start-ups, who need it for their development, thus encouraging market progress[67]. While the framework has found the support of researchers we spoke to – unlike the voluntary system, compliance mechanisms are assured through the imposition of penalties. Without this, contractual relations will be prioritised over market power, which is unlikely to change the current dynamics of the system. However, they cautioned that it will only work if there is transparency in terms of informing the farmers about the data being collected and its future use.

As stated previously, businesses participate reluctantly in such frameworks and often actively reject the imposition of these provisions upon them.

Type of regulation	Mandatory provision	Justifications	Response by businesses
Digital advertising	Australia's Consumer Commission recommended for big tech companies like Google to share users' anonymised personal data with advertisers	Foster competition	Google threatened to pull out its search engine services from Australia; *policy under discussion*
Non personal data sharing	India's NPD Framework mandates for 'data businesses' to share non personal data collected with the Government	Innovation of digital economy	Google, Amazon raised concerns about the data sharing; possible threat to IP rights of companies; FDI might be limited in India; *policy under discussion*
Intermediary rules	Under India's IT Rules intermediaries (WhatsApp, Twitter) must share information about 'originators' of 'unlawful' messages	Protection of sovereignty, integrity of India	WhatsApp, Google approached Delhi – rules violate privacy, freedom of speech of citizens – refusing to comply; *rules came into force in May 2021*
Data localisation	Foreign companies are required to store data of Indian users within the territory	Protection of privacy of Indian users	Amazon, Google raised concerns – increase in compliance costs, forced change in business models, planned investments affected; *policy under discussion*

Figure 8: Private sector response to mandatory provisions

Image source: https://www.freepik.com/free-photo/young-green-corn-growing-field-background_17830932.htm#query=farmer&position=15

[67] "Enhancing access to and sharing of data", OECD Library (November 26, 2019) retrieved on August 3, 2021 from https://www.oecd-ilibrary.org/science-and-technology/enhancing-access-to-and-sharing-of-data_276aaca8-en

Conversations with small agricultural businesses and start-ups uncovered concerns with the mandatory data sharing system pushed by the NPDR. The vague definition of 'data businesses' leaves them wondering whether they fall in the category. This gives them little time to prepare for the increasing compliance costs that accompany the classification (they will have to set up the technology to maintain and update their metadata). The definition of non-personal data in the NPDR is very broad – start-ups, in particular, may not be able to accurately distinguish between personal and non-personal data, especially when it comes to mixed datasets. Absence of personal data protection legislation adds to these woes[68].

The regulations provide little incentive for them to share their data – they are only allowed to charge a 'nominal amount' for the processing stage (the reasoning being that data collection is already a part of their business process, and therefore the expenditure for this is pre-decided). This might negatively impact the current data sharing system in which companies voluntarily share internal data for public purposes. For example, Facebook's Disaster Maps initiative shares public emergency data (like information about those impacted during natural disasters) with trusted partners in an aggregated form. Mastercard's 'Center for Inclusive Growth' shares donation insights reports containing anonymised and aggregated transaction data with small companies for them to learn more about the trends in philanthropic donations[69]. Apart from this, there is no provision in the NPDR about compensation to businesses for the losses they might face as a result of sharing data. While it makes mention of protection of IP rights, particularly of trade secrets, it hardly delineates how this will work on the ground.

NPDR views data from an antitrust perspective, which may not be necessarily true in all cases. The mandatory framework assumes public welfare by moving data away from the control of select private entities. However, it does not address questions of stakeholder incentives and acknowledge the contribution of the private sector in innovation. 'Dominance' is assumed to be a synonym for abuse of power, which is an inaccurate assumption. This can put big businesses, which make significant contributions to the sector, at risk. Apart from this, the NPDR only targets businesses in the offline world. Platforms in the online spaces can also cause harm to users and make problematic antitrust decisions. This issue of deciding the standards for 'dominance' for internet businesses has been left unaddressed[70]. This information asymmetry must be cleared by the Committee before legislating on non-personal data, lest it face staunch resistance from the private sector.

Image source: https://www.freepik.com/premium-photo/indian-farming-technique_5010311.htm#page=3&query=farmer&position=30

[68] "Non-Personal Data Regulatory Framework: Community recommendations from India's startup and investor ecosystem.", HasGeek, retrieved on September 2, 2021 from https://drive.google.com/file/d/1Z8OGQ88_L19kyyTxIQ3WRPxbDpqZKc8G/view

[69] "The Global Commons of Data", Shkabatur J., Stanford Review (2019) retrieved on August 3, 2021 from https://law.stanford.edu/wp-content/uploads/2019/09/Shkabatur_Global-Commons_20190830-1.pdf

[70] "Five Key Concerns With India's Non-Personal Data Report", Medianama (July 22, 2020) retrieved on August 3, 2021 from https://www.medianama.com/2020/07/223-five-key-concerns-with-indias-non-personal-data-report/

Mandatory data sharing does not account for interests of farmer communities

- The term 'communities' as defined in the NPDR is broad and vague – can lead to wrongful interpretation, leading to harassment and regulatory capture.
- The difference between community non-personal data, and public and private non-personal data is unclear.
- The power relationship between the data trustee and the community is not defined – not the most appropriate representative of the community.
- Issued relating to the right to privacy not appropriately addressed in the regulations – drafted in the backdrop of an absence of a legislation on the same.
- 'Data businesses' (which can include community-based organizations) that cross the regulatory threshold will be mandatorily required to share data with the Government - compensated only for data processing and not data collection.

Private sector not sufficiently incentivised to participate in the data sharing process

- Lack of clarity in the NPDR about which entities will contribute data to metadata directories and creation of HVDs.
- Start ups are unclear about what part of the data constitutes personal data and non-personal data.
- Most startups are unaware of NPD framework and its possible implications because the entire process of drafting has been made inaccessible from the start.
- Data (particularly metadata, which has to be shared) can reveal company strategies which can affect competition between them.
- Increase in compliance costs can be a deterrence for innovation or even the entry of start ups into the market.

Figure 9: Drawbacks of a mandatory data sharing framework

Image source: https://www.freepik.com/premium-photo/potato-plantation-field-with-loosened-soil-loose-crushed-moist-soil-after-cultivating_17076617.htm#position=7

6.2.

Voluntary data sharing frameworks

Research into the impacts of the mandatory data sharing framework led to findings of alternative modes of data sharing that have been adopted globally. These include free form voluntary sharing, ecosystem-enabled voluntary sharing, sector-specific mandates, and purpose-specific mandates. Our research showed that agricultural sectors in major economies have chosen to go with an ecosystem-enabled voluntary framework. Such a framework aims to promote altruistic sharing and innovation through the establishment of a trusted ecosystem, while reducing regulatory and capacity burdens across stakeholders. An ecosystem-enabled voluntary framework follows a bottom-up approach while drafting policies, thus ensuring that the provisions are sensitive to the needs of the stakeholders. The focus of policy in such a framework is on mitigating existing challenges in data sharing and investing in infrastructural, legal and technical building blocks that foster a trusted network for voluntary sharing. Incentives for participation are also prioritised in such a framework, with this approach allowing stakeholders a chance to divide the burden of conceptualising and testing, with the resultant policy being reflective of varied needs and concerns across the board. Such a framework is premised on a strong foundation of robust standards for ecosystem building, resolute protocols for data sharing, and reliable data sharing infrastructures. In building such an ecosystem, opportunities for economic incentives are thus created organically. With a reliable and stable ecosystem in place, companies are left to share/innovate voluntarily, based on their individual capacity. It is also helpful to build the ecosystem based voluntary data sharing system through the lens of a particular sector; this helps contextualise data sharing and makes it more robust. Voluntary frameworks adopted by global jurisdictions demonstrate the importance of setting up a single-point, specialised regulator that can help prioritise and protect the interests of stakeholders. The framework ensures continuous, active participation by the community members, which automatically translates into their work – being updated on the issues of data rights can help them make more calculated decisions while sharing their information with third parties.

Image source (top): https://www.freepik.com/free-photo/young-green-corn-growing-field-background_17830932.htm#query=farmer&position=15
Image source (bottom right): https://www.freepik.com/premium-photo/indian-farmer-labour-drip-irrigation-pipe-assemble-agriculture-field-rural-scene_14721729.htm?query=farmer

6.2.1. International best practices

Set out below are instances of adoption of such a model by agricultural sectors in jurisdictions worldwide, with the successful adoption a result of acceptance by both the public and private sector.

A. United States

The United States' "Privacy and Security Principles for Farm Data" (Principles for Farm Data) was established by the American Farm Bureau in 2014. The framework received the support of the government and over 2015 and 2016, the House Agriculture Committee began hearings on data practices in agriculture, and then set up the Agricultural Data Coalition, a neutral, secure farmer-centric data repository.
The objective of the agreement is to ensure that access and use of farm data be done only with the explicit consent of the farmer. The code attributes ownership of the data to the data subjects – the farmers – which is transferable or alienable through bilateral contracts. This idea has been reiterated in the EU's agricultural data sharing policy as well. It has been argued that this represents a departure from the inalienability right argument postulated by the General Data Protection Regulation (GDPR). However, the GDPR does allow data to be shared if the data subject consents. This principle is echoed in the above policies, allowing for the sharing of data for economic development, while prioritising the interests of farmers. Data portability is considered to be a privacy-related issue, rather than a competition issue. This right, however, only extends to non-anonymised and non-aggregated data. The policy also pushes for complementary protections like the principle of purpose limitation and data deletion. The authors of the agreement – farmers and organisations representing farmers' interests – set up the Ag Data Transparent, an evaluator to verify compliance with these provisions by businesses. Farmers were confident about sharing data with companies that had received the affirmation of this regulator. Large businesses like John Deere recognised its importance and updated their policies in return for the stamp of trust. Other businesses drew inspiration and as of 2020 approximately 20 agricultural organisations are registered with the regulator.

B. European Union

The European Union Code of Conduct on Agricultural Data Sharing by Contractual Agreement (EU Code of Conduct) was launched in 2018 by a coalition of associations in the EU agri-food chain. These non-personal data regulations immediately followed the release of the EU's GDPR on personal data. It promotes digitisation of agriculture, while taking into account the interests of agri-cooperatives and agri-businesses. The data of the farmer can only be shared after confirming their consent, preferably through a written agreement or contract. The policy gives ownership rights for data generated during farming operations to the originators of the data – the farmers. They have a right to benefit from and/or be compensated when this data is further shared with third parties. However, if the data is collected using external machinery, then the operator of that machinery is considered the originator. This can have implications, particularly for non-personal data such as weather and soil data, which are collected and analysed by businesses using specialised machinery developed by them. The right to port data is vested within the data subjects, unless otherwise agreed by the subject and the third party. Data transfers can be done when it is technically feasible, and in a manner that does not harm the sensitive personal data of the subjects. The right to

deletion and removal is reflected in the right to be forgotten as propounded by the GDPR. Similar to the US, this right does not extend to aggregated data that has been processed.

C. New Zealand

The New Zealand Farm Data Code of Practice, Standards and Accreditation (NZ Data Farm Code) was set up in 2014 and drew inspiration from the US policy. It is a set of guidelines "enabling the sharing of data within the agriculture industry". Businesses are mandated to inform data subjects about the rights they have over the data, the rules involved in the sharing process, details about the security in place to protect the information and the jurisdiction in which it is stored. Data must only be used for the purposes agreed upon by the subjects. Businesses that comply with the standards of the code are required to display a 'Code of Practice' mark on their website and other relevant documents. They must submit proof of compliance to the NZ Farm Code regulator – if approved, they receive a licensed trade mark. Approximately 45 businesses have confirmed compliance with this code as of 2020.

D. Australia

The Australian Farm Data Code was released by the National Farmers Federation in February 2020. In 2017, the Australian government released the "Productivity Commission Report on Data Availability and Use in Agriculture", which began the discussion on this topic. Similar to the NZ Code, the objective is to promote digitisation in the farm sector and build the trust of farmers while collecting and sharing data. The Code applies to entities that either have a direct commercial relationship with the farmer or collect and manage data in connection with this relationship. It divides data into three distinct categories – public data, farm data (where the originator is the farmer) and private data. Some of the principles laid down in the code include fair and equitable sharing and use of data, ability of the farmer to control and access farm data, portability of the information and security of the data.

Image source: https://www.freepik.com/premium-photo/grazing-sheep-pasture-field-sheep-are-black-brown_18055456.htm#page=1&query=farmer%20new%20zealand&position=15

6.2.2. Points of concern

While the voluntary data sharing framework has seen a wide adoption rate, there are a few points of concern that need to be kept in mind by policy makers adopting this route. First, participation of businesses in the framework is not widespread and only a few businesses have agreed to comply with these codes. As these codes are voluntary, businesses are not obliged to adopt them. Since contractual relations overtake market forces, businesses can continue to use their network effect to widen user database, even without getting the certificate from regulators. Second, there must be a technically feasible environment to ensure that these codes are successfully implemented. This requires governments to make investments in the digital economy, which might not always be the case. Additionally, data interoperability might be a problem – businesses will have their own technical standards while sharing data, and unless APIs are made uniform, this process will be difficult. Finally, the codes are drafted by the coming together of the farmer community. This entails the overlapping of different interests, which might conflict with one another – drafters may find it difficult to accommodate all of them in the policy. Majority representation, while important, might lead to the exclusion of small but important opinions.

Despite the concerns, it is clear that an ecosystem-enabled voluntary data sharing mechanism works best and has a high possibility of addressing issues and loopholes in the mandatory system, by allowing for resolution of issues effectively in both a top-down and bottom-up manner. It will also achieve the NPDR's objective of maximizing public value and dealing with antitrust issues.

Resolution of issues (top-down perspective)	Resolution of issues (bottom-up perspective)
• Businesses incentivised to share data with Government – compensation accounted for (for collection and processing of datasets), IP rights given due protection • Data markets can be set up – allows for transparency, checks imbalance of power between big businesses and start ups during data sharing process • All businesses encouraged to participate, irrespective of data drivenness – equal enabler system • Clear protocols for safe sharing of data established – privacy legislations to be complied with, stakeholders incentivised to participate • Compliance costs reduced – opportunity for businesses to redirect finances for innovation purposes • Businesses are able to strike a balance between innovation and public value	• Farmer community consulted while deciding on policy – sector specific issues included in the drafting stage • Data subjects rights protected – businesses mandated to adhere to privacy laws – issues of consent, data anonymization automatically addressed • IP rights of farmers prioritised – transparent system allows for checks on imbalance of power while entering into contracts; farmers (either directly or via data stewards) can claim protections for farmers rights • Incentives to participate addressed – Farmer community either given monetary compensation or agri-advisory that will help improve incomes • Community incentivised to actively participate in data sharing process; digital literacy trainings can be set up to introduce them to privacy rights of data subjects

Figure 10: How a ecosystem-enabled voluntary framework can resolve issues

07

Recommendations

The current NPD policy classifies data businesses horizontally[71]. This approach, however, does not account for sector-specific needs and overlooks issues like incentives required for stakeholder participation. ns. Sectors vary in terms of their digitisation and impact on the economy – a prime example being the low levels of digitisation in the agricultural sector when compared to the mobility sector[72]. The NPDR also regards the act of data sharing as an end point in itself, and does not consider broader ecosystem level questions.

Therefore, our approach focuses on a process of data sharing that is primarily driven by sectoral stakeholders, pivoted on the factors that encourage voluntary participation in the process. In order to do this, the approach must go through four stages:

01

Setting up of a sectoral body to ensure the implementation of the framework.

02

Build technical capacity through key first efforts.

03

Provide incentives for stakeholders that will help add to public value.

04

Test NPD sharing through sandboxes and pilots to further innovation.

With this in mind, we make the following six recommendations

1. An inclusive consultative method must be used while developing data sharing policies

Government policies in general, and the data sharing policy of the NPDR in particular, have been drafted without adequate consultation with the primary stakeholders. In this case, it includes farmer communities, small agricultural businesses and start-ups. Representation has instead been made through non-governmental organizations that try to reflect the voices of stakeholders but submissions are not public, and there is no way to ascertain if farmer interests have been represented at all. Just like in the BT-Brinjal consultation in 2010, Governments must send representatives on the ground to explain the policy and get the feedback of these stakeholders.

[71] "Govt invites suggestions on non-personal data framework by July 19", Alawadhi N., Business Standard (July 12 2020) retrieved on June 28 2021 from https://www.business-standard.com/article/economy-policy/govt-invites-suggestions-on-non-personal-data-framework-by-july-19-120071200914_1.html

[72] The Pandemic will accelerate the digitization of the Indian automotive industry: Auto Experts at OLX Auto Talk 3.0", The Print (June 17 2021) retrieved on June 28 2021 from https://theprint.in/ani-press-releases/the-pandemic-will-accelerate-the-digitization-of-the-indian-automotive-industry-auto-experts-at-olx-auto-talk-3-0/679850/

2. Incentives must be included in the policy to encourage stakeholders to participate

Data sharing frameworks are a relatively new and evolving concept in India. More particularly, mandatory data sharing leaves stakeholders with little choice but to participate in the system. There may be disincentives to participate in mandatory regimes which may result in non-compliance, retreat from the market, hampering innovation, and mistrust of the Government. It is, thus, important that incentives such as reciprocity, collaboration, compensation for collecting and sharing data, investments by the Government in the agricultural sector, improved technology for interoperability and assured standards for data protection be included as a prerequisite to any data sharing recommendation, especially if the government chooses to go down mandatory pathways.

3. Adopt an evidence-based approach to examine data sharing frameworks

A common thread that was found between global jurisdictions that identified data sharing frameworks was that they tried different means of policies, consultative research and testing. In a similar manner, instead of deciding on a mandatory data sharing framework, which has no evidence from the ground, the Government must take time to allow stakeholders to experience different data sharing mechanisms and consider the feedback through pilots and sandboxes, prior to implementation. This provision of choices and tests gives flexibility to stakeholders and builds buy-in to participate in a data sharing ecosystem that engenders trust.

4. A sector-specific policy might be more effective than a generalised data sharing policy

The benefit of a sector-specific policy is that it can be tailored to the specific needs of the various stakeholders in the sector; examining the impact of the NPDR on agriculture has made this abundantly clear. Take for instance the National Health Data Management (NDHM) policy which focuses only on the needs of the health sector and the data sharing instances that take place there. A general framework like the NPDR might not be able to cover these nuances, which will lead to the missing out of important factors such as different incentives needed for stakeholders to be encouraged to participate.

5. Establish trust-based mechanisms for data stewards to enhance community-level trust in data sharing

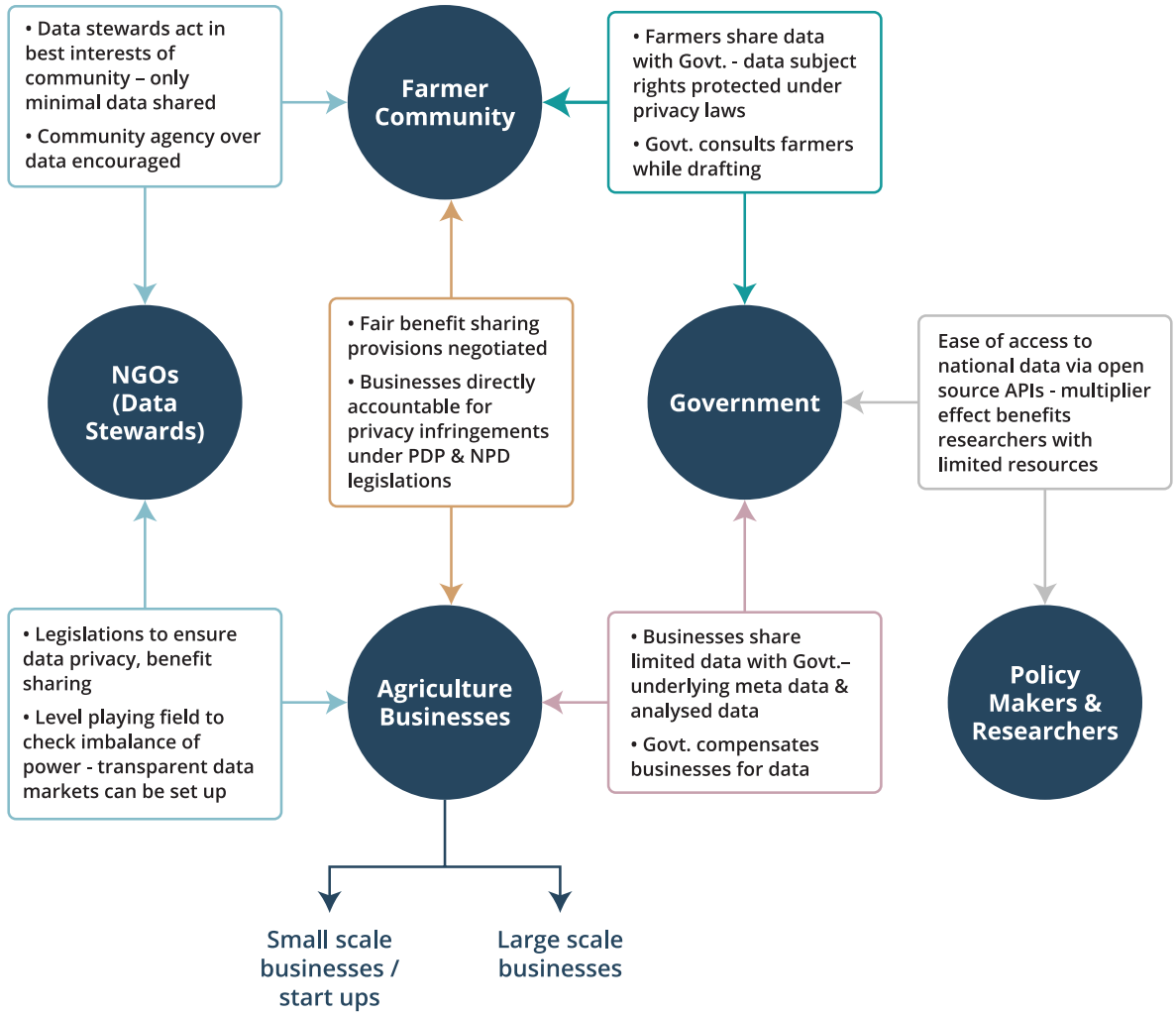
To enhance trust within communities, and incentivize stakeholders to participate in voluntary data sharing, data stewardship models need to be encouraged. The NPDR already mentions data trustees, but their role is unclear, and are likely to see capture from other interested parties. There are other examples like consent managers which can be enhanced to become stewards, that work to safeguard the interests of communities and ensure they are able to minimise harms and draw value from their data. Community data stewards can enable communities to draw the greatest value from their data. Incentive measures could include translating policies to regional / local languages as well as creating digital learning opportunities for farmers.

6. Adopt an ecosystem data sharing approach

We suggest that the Committee of Experts on NPD rethinks its recommendation on mandatory sharing to a more ecosystem-led voluntary approach that brings together different stakeholders and is structured around incentives instead of mandates. An ecosystem approach involves investments, both by the government and the private sector, in the physical, technological and human infrastructure required for sharing, and to co-create solutions for public interest questions. This approach ensures greater compliance and ownership from the private sector, and also makes the government an active player in data sharing. It further enhances the role of communities and community based organisations that have the opportunity to carve out their roles in the ecosystem.

These recommendations will help implement an ecosystem-enabled voluntary framework for data sharing that adequately accounts for stakeholder interests and addresses the drawbacks in the existing and NPDR recommended mandatory data sharing system.

Image source: https://www.freepik.com/premium-photo/old-man-was-watching-field_15942442.htm



- **Methodology** for data collection vague; **data protection** policies not specified – **data subjects rights** unknown; no support to protect **IP rights** of farmers
- Analysed data shared on Government website – **inaccessible** to farmer community; **indirect impact** felt on community via new policies/ improvements in existing policies ; farmers **excluded** from policy making

- **Quality of datasets** maintained by Ministry questioned by businesses; **complex procedure** for accessing data discouraging; businesses **excluded** from policy making
- **Limited data** shared with Government, businesses' IP interests given priority; data sharing done at **high costs**, burden on Government

- **Quality of Government** datasets questionable; **complex procedure** for accessing data discouraging for researchers; results of research not shared with farmer community (primary stakeholders)

- Data stewards collect data **haphazardly** – may not be aware of data subjects privacy rights; burden to update consent, understand nuances of business contracts, make ethical decisions in farmers interests
- NGOs share data as stewards – **imbalance of power** – NGOs forced to negotiate; may not have resources to understand complexities of **contracts** with businesses

- **Imbalance of power** between businesses and farmers – unfair provisions in contracts (example – technological lock-ins, excess data sharing, non-purpose specific use of data); **IP rights** of farmers susceptible to infringement
- Agricultural advisory given by businesses to farmers at **high costs**; rights of data subjects given **lesser priority** over business interests

Figure 11: Indian agri data sharing framework built on an ecosystem-enabled voluntary approach

08

Conclusion

The NPDR is an ambitious effort at creating a framework for the governance of data sharing of NPD in India. However, the top-down approach adopted in its preparation has led to the exclusion of key stakeholder inputs, and results in a framework that does not adequately address the existing issues in the data sharing ecosystem. Conversations with farmer communities, agri businesses and sectoral experts undertaken as part of the research indicate a sense of distrust among key stakeholders towards the data sharing ecosystem proposed by the NPDR. A review of data sharing systems implemented in agriculture sectors across various jurisdictions highlights that an ecosystem-enabled voluntary approach to data sharing is most effective. However, in order to implement such a system there are open questions that require further research and study. These include how the gap in digital literacy is to be addressed, which is linked to the question of how best to ensure that benefits of data sharing reach farmers, and identifying suitable standards and taxonomy for interoperability specific to the Indian agriculture sector.

Image source: https://www.freepik.com/premium-photo/young-indian-farmer-standing-cotton-agriculture-field_18064789.htm#&position=22

Annexures

Annexure A

Case study: Possible impact of the NPDR on community-level organisations – Digital Green and Jaljeevika

To include a community-level perspective of the issue, in addition to the top-down analysis of policies, we conducted a case study of two community level organisations - Digital Green and Jaljeevika.[73]

CASE STUDY 1 - DIGITAL GREEN

Overview of Digital Green

Digital Green (DG)[74] was set up as a global development organization that specialised in providing agricultural-based technology advisory to farmer communities. They began working in India, and eventually scaled up to global jurisdictions like Ethiopia, South Asia, Latin America and Sub-Saharan Africa. Initially, DG used technology as an entry point for their work - they assisted[75] farmers in making videos on relevant agronomic issues, which were disseminated to other members of the community. DG partnered with researchers, practitioners and farmers to understand challenges that needed to be discussed, and this went on to become the topics for the videos. These included issues such as agricultural practices and nutrition behaviours that could be adopted by farmers to improve their livelihood.

In order to scale up their knowledge creation and exchange mechanism, DG set up a database - Connect Online Connect Offline (CoCo) - containing personal and non-personal data of farmers. Moving away from a one sided relationship of collecting, analysing and sharing data with farmers, they made the platform freely accessible and updated their findings based on regular engagements with the community. The analytics dashboard helps farmers visualise their data, helping them gain insights that form the basis of critical agricultural decisions.

With the development of Government databases for the agricultural sector, the question of interoperability came to the forefront. Data exchanges between public bodies and organizations like DG were needed to get an all-round understanding of the sector and to establish feedback mechanisms between the two. These developments came with questions concerning data privacy of farmers, particularly that of confirmation of consent in the sharing process.

[73] Jaljeevika website available at https://www.jaljeevika.org/
[74] https://www.digitalgreen.org/
[75] https://www.digitalgreen.org/videos/

Methodology

For the purposes of this report, we used three methods to understand the working of DG:

Interviews with internal members of DG

We spoke to the Director of the programme strategy at DG and and a Platform Architect, who helped us understand their work from an administrative level.

Review of internal documents

Particularly those relating to FarmStack (FS), a consent manager set up by DG. This gave us an overview into the underlying principles of the platform, the details of the technical architecture and the usage control policy adopted to protect the privacy of its users.

Interview with members from the Food and Agricultural Organization (FAO)

As beneficiaries of DG's work, we were able to get a first-hand understanding of DG's work at the community level

Challenges in setting up agricultural databases

Agricultural datasets set up by businesses that contain personal and non-personal data of farmers are generally inaccessible. Proprietary models of data collection adopted by the private sector exclude farmers from getting access to the inferred data. This makes interoperability difficult for the farmer - for example, shifting from one company to another is almost impossible because they cannot get access to historical data.

The need to set up public agricultural datasets has been argued for by international organizations like Global Open Data for Agriculture and Nutrition (GODAN)[76]. Making data transparent and accessible will be beneficial[77] for all actors in the agricultural value chain. The information can also be re-used for larger public purposes such as optimization of agricultural practices, stimulation of sectoral finance and implementation of policies to enhance sectoral efficiency.

The Government of India's intervention into the agricultural sector to promote the sharing of data for public purposes has been criticized by civil society organizations. For instance, in 2020, details about the Agristack platform came out in response to a Parliamentary question. The platform is supposed to be a centralized database that contains personal, sensitive personal and non-personal data of farmers from across the country. According to this explainer[78], there is a possibility that banks and insurance companies might get access to this information, particularly financial information, on the basis of which they will make crucial decisions like the granting or denial of loans or the amount of premium given. There is no documented evidence of a pilot having been carried out, or of a consultative process in assessing the impact of this platform before its rollout. Not only does this put farmers at risk, it also infringes upon their agency over the data and their right to privacy.

[76] https://www.godan.info/pages/open-data
[77] https://www.godan.info/working-groups/agriculture-open-data-package-working-group
[78] "The AgriStack: A Primer", Internet Freedom Foundation, retrieved on August 2, 2021 from https://internetfreedom.in/the-agristack-a-primer/

A study done by DG in Ethiopia showed that farmers were not keen on sharing with organizations due to unresolved questions relating to data security and ownership, which culminated into a lack of trust in the system. User agency over data was another point of conflict that emerged from the study - with developments in technology, farmers are moving further away from their data, and with this, their right to decide how it is used and with whom it can be shared.

With this background, DG began to base its work on the increased control over data and sharing decisions by the farmer.

DG's solution to improved user agency - FarmStack

Developed as a consent management system, FarmStack (FS) is an open source protocol that functions as a peer-to-peer system to ensure the secure transfer of data between organizations and farmers. Rather than establishing itself as a platform, FS made the conscious choice to work as a protocol. Platforms are centralized structures that work on the principle of "winner takes all" (the more the number of users that register with them, the higher is the amount of data collected). Protocols on the other hand are decentralized frameworks that allow any user to build their functioning off of it. This helps them create network effects, with a focus on developing public good.

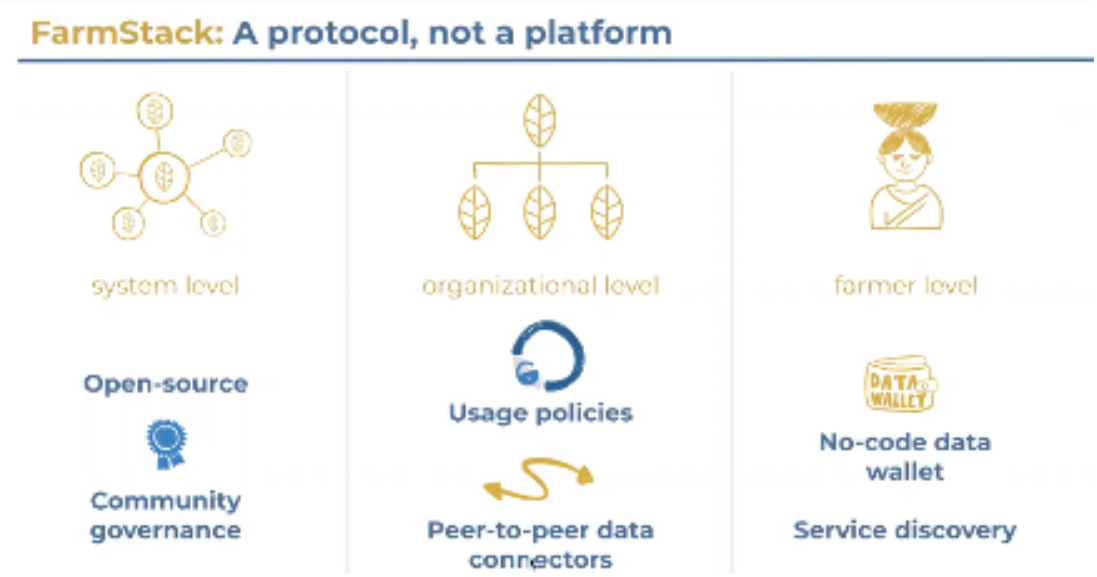


Image 1: Elements that make FarmStack a protocol

As an open source protocol, FS gives their users (farmers) greater control over their data through the use of data wallets and usage policies. Data wallets contain the personal and non-personal data of farmers, which is inputted by them over time. They function as a personal data management system (PDS) which allows users to store data in them, giving them the agency to decide when and with whom they want to share the data. PDS' are still not regulated in India, which means that the PDP and NPDR regulations will be applicable depending on whether the data stored in the wallets is personal or non-personal.

How does FS work?

A farmer gets a contract that involves an opportunity to share data with another organization who agrees to sell the produce in return for a profit. The organization (a data steward) shares this contract with the farmer and explains the terms and conditions to them. If the farmer consents to sharing their data, they confirm the same on FS' application.

FS then shares the relevant data with the receiving organization through their peer to peer connectors in a privacy preserving manner. Being data blind, FS cannot see the information being transferred nor do they store it in their database. A farmer merely has to update the FS application if they want to stop sharing their data with the other organization.

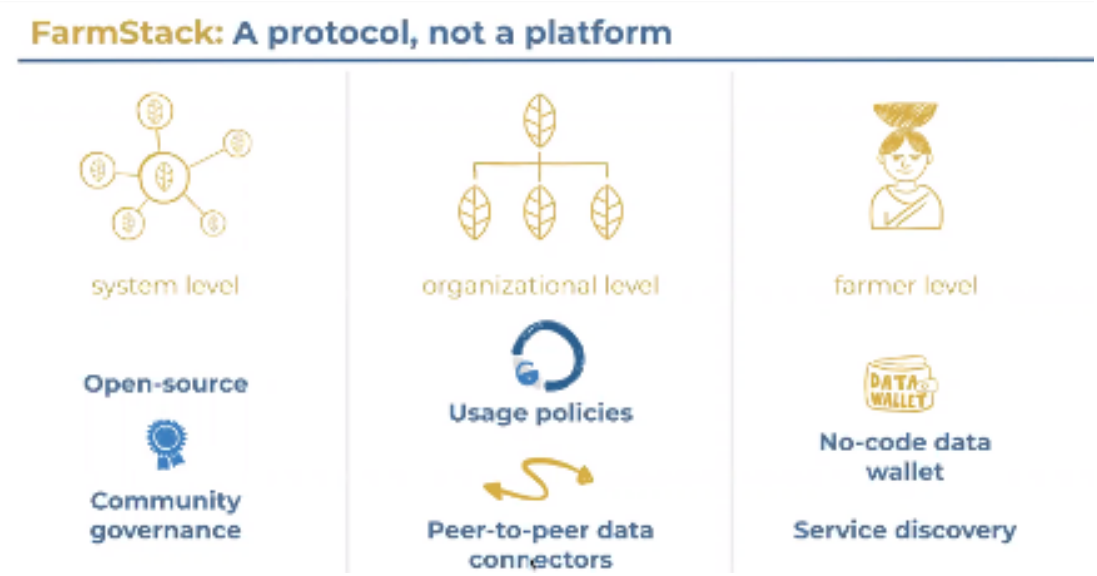


Image 2: The Farmstack Data Exchange

Organizations develop usage policies which are a reflection of the principle of 'purpose limitation' - data can only be shared for the purposes for which the farmer has given their consent. Currently, FS works on an organization to organization level - they plan to shift to direct interactions with farmers in the future. FS is liable for harms that occur to the data only to the extent that it violates the usage policies. For instance, if the recipient organization uses the data for purposes other than what was agreed upon, FS will reject the sharing of the data and it will be returned to the original organization.

FS does not go into the question of the quality of data being transferred, as this is the responsibility of the data steward. They work on the assumption that the data being transferred is correct - quality information ensures quality services. The protocol follows a global standard of interoperability while sharing data which has been set by the Fraunhofer Institute[79] in Germany. This allows organizations to share data irrespective of where it is stored (whether in their own servers, in a central cloud etc) This is beneficial for the smooth flow of data during exchanges between organizations or between the private sector and Government.

Currently in its pilot stage, FS is being used by farmers in Ethiopia to make transactions. The protocol is working to improve itself on the basis of this feedback, after which it will be introduced into the Indian agricultural sector. It is currently being used at a business-to-business level and is not optimized for direct interactions with farmers. DG hopes to eventually expand the use of FS to the latter stage in the coming years.

[79] https://internationaldataspaces.org/

Critical analysis of FS as a consent management system

The Data Empowerment and Protection Architecture Framework (DEPA)[80]. The framework defines consent managers as conduits "for encrypted data flows". Consent management systems have been adopted as 'account aggregators' by the Reserve Bank of India (RBI) to enable secure sharing of financial data. FS' work also overlaps with the definition of data fiduciaries[81] in the Personal Data Protection Bill, 2018. Since both the policies are under discussion, it is to be seen how exactly their provisions will impact the working of FS. The specific adoption of consent managers by the RBI gives rise to the hunch that consent managers may be regulated[82] at a sectoral level, aside from the general policy of being mandated to register with the Data Protection Authority.

There is also little clarity on their business model. If they are simply neutral structures that transfer data of individuals in an encrypted manner, where do they get their profits from? Consider the situation where they are given monetary compensation for their services from the profits made out of the data exchanges they helped initiate. Wouldn't that give rise to the possibility of them having vested interests in these exchanges, thus pushing organizations or farmers towards more profitable relationships? Unless regulations for consent managers are established, the reluctance to engage with them would be understandable. A point of interest here would be direction 5 (e) of the Directions[83] for Account Aggregators issued by the RBI - it prohibits them from engaging in any business apart from that of aggregation. It will be interesting to see how this will apply to FS, owned and operated by DG, an organization that is simultaneously engaged in other community-based businesses apart from this.

Consent managers do not necessarily take the burden off of data stewards. For instance, it has to continue its job of confirming consent of data subjects (farmers) by explaining to them the terms and conditions of the contract. FS offered a possibility where farmers' consent fatigue can be addressed, while simultaneously reducing the burden on data stewards to collect and update consent for every new contract. A one-time consent could possibly be given per season, which would lead to the automatic transfer of farmers' data for any new contract. This could be problematic because it may end up in a situation where farmers may consent to contracts that they did not actually want to enter into in the first place.

Being data blind, FS is not responsible for the quality of the data being transferred. This means that the data stewards responsibility of correcting and updating information to ensure its quality still rests on its shoulders.

Adopting a consent management model might also increase the business costs of data stewards. For example, they will be required to monetarily compensate FS for its services. This argument is backed by the DEPA framework which says that consent managers can charge a nominal fee for their services for example, through a subscription model.

Through this analysis, we conclude that while FS' intention and functioning might benefit the user agency and privacy of farmers, data stewards might not be inclined to adopt them. The only way forward is for future regulations to give more visibility on their functioning and accountability mechanisms in lieu of data sharing relationships.

[80] "Data Empowerment And Protection Architecture", NITI Aayog, accessed on September 19, 2021 from https://niti.gov.in/sites/default/files/2020-09/DEPA-Book_0.pdf
[81] "ANALYSIS OF THE NEW DATA PROTECTION LAW PROPOSED IN INDIA", Nishith Desai Associates, accessed on September 19, 2021 from http://www.nishithdesai.com/fileadmin/user_upload/pdfs/NDA%20Hotline/Analysis_of_the_new_Data_Protection_Law_Dec2419.pdf
[82] "Regulating Consent Managers in India: Towards Transparency and Trust in the Digital Economy", Basu S. and Sonkar S., Oxford Business Law Blog (April 1, 2020), retrieved on September 19, 2021 from https://www.law.ox.ac.uk/business-law-blog/blog/2020/04/regulating-consent-managers-india-towards-transparency-and-trust
[83] "Directions regarding Registration and Operations of NBFC - Account Aggregators under section 45-IA of the Reserve Bank of India Act, 1934", Reserve Bank of India, accessed on September 19, 2021 from https://www.rbi.org.in/Scripts/bs_viewcontent.aspx?Id=3142

How will the Non-Personal Data Regulations (NDPR) affect the working of DG?

At the outset, a clear distinction will have to be made by the NPDR between the data collected by DG through CoCo and the data shared through FS - the latter is a consent manager and cannot be mandated to keep records of data and share it with the Government.

The NDPR has created new institutions, each having a separate role in the process of collecting and managing non-personal data. The objective of all of them is similar - to share data in a manner where the interests of the community are maximized, while also minimizing any harms that may occur to them.

Prima facie, this is what DG does. It prioritises the interests of its users by giving them agency over their data, while also ensuring that it is only used for the purposes agreed upon. This might lead to DG being classified as a 'data trustee' under the NDPR. 'Data trustees' are institutions that create, maintain and data, all in the interests of the community. They can either be a Government body, a not-for-profit organization or can be formed by the coming together of members of the community. Data trustees have a duty of care over the data that they control and are obligated to set up grievance redressal mechanisms in case any harm occurs to the user's data. For example, harms that may occur in case of re-identification of an individual from the data shared. These trustees can identify opportunities to combine data from multiple businesses for the benefit of the community. The NPDR unequivocally declares all data trustees to be data businesses. They can charge a nominal fee for performing the function of data processing (and not data collection, as the NPDR assumes that it is a part of their business). This means that if DG crosses the threshold set by the NPDR (which is yet to be decided) it will have to mandatorily share farmers data with the Government. Confirming consent before collecting data has always been important. With the possible introduction of the mandatory data sharing framework by the NPDR, data will have to be shared irrespective of consent for public good purposes. Therefore, DG's responsibility of getting the consent of the individual farmer at the initial stages of collecting data becomes crucial.

DG may also be classified as a 'data custodian' - an entity that has a direct relationship with the data principal and collects, stores and manages this dataset. It can either be a Government body or a private organization which has a duty of care towards the data collected. Custodians are obligated to share non-personal data if requests for this information abide by the standards set in the NPDR. This data must be shared in an anonymized form, with due protections taken to protect the user from any harm. Data custodians too have been declared to be data businesses, which means that if DG is classified as a data custodian, it will have to mandatorily share data of farmers with the Government.

Image source: https://www.freepik.com/premium-photo/beautiful-tea-plantations-south-asia_18056884.htm#page=4&query=farmer%20india&position=7

CASE STUDY 2 - JALJEEVIKA

Overview of Jaljeevika

Jaljeevika (JJ) is a not-for-profit organisation that works with farmers in the aquatic sector in Andhra Pradesh, Bihar, Jharkhand, Madhya Pradesh, Maharashtra and Odisha. It is a community-based organisation that builds technological solutions from the ground up by integrating knowledge at the local, regional and national levels. Its work involves facilitating low-cost technological transfers, building effective institutional governance models and encouraging the development of an innovative and resilient blue economy. With this, it hopes to enhance the livelihoods of farmers and provide them with food and nutritional security.

JJ collects data from multiple stakeholders, some of whom include[84]:

Farmers

From farmers it collects personal data such as the name of the farmer, village in which they live, financial data and non-personal data such as crop yield or output of the farmer at the end of each season.

Other members from the community, like those from Self-Help Groups

For example, they provide information relating to pond ownership.

Farmer Producer Organisations (FPOs)

Personal data about members of the organisation is collected, including details of inputs made to grow crops and total crop yield for each season.

Community resource persons or agents

They are selected from amongst the residents and trained by JJ to collect personal and non-personal data from farmers assigned to them. JJ believes that if the data collection process is done by a familiar and trusted face, farmers will be more comfortable in sharing their data. These agents are also trained to explain the terms and conditions of any new contract that the farmer might be interested in registering for, whether with a business or the government.

Vendors

JJ collects their details, including their names, contact information and products they sell.

Image source: https://www.pexels.com/photo/green-leaf-closeup-photography-1592119/

[84] Jaljeevika Strategy Plan (2020-2050) - https://secureservercdn.net/160.153.137.14/yjd.b40.myftpupload.com/wp-content/uploads/2020/04/Strategy-Document-1.pdf

The data collected is then processed by JJ to be stored in their database. This is done using three platforms[85]:

Epicollect5 and Kobo

An open source mobile data gathering platform on which raw data from farmers is collected. They provide both web and mobile applications for the generation of forms (questionnaires) and are freely hosted project websites that enable data collection. Their IoT-based application which is currently being developed is responsible for collecting non-personal data from the farmers. Limited financial data is collected on this application, until such time as they come up with stronger encryption mechanisms to ensure security.

Google Sheets

This data is then inputed into excel sheets and cleaned for quality purposes.

Google Data Studio

Information from the data is unlocked using Google analytics and represented in the form of interactive dashboards. Accessing data in this manner makes it easier to take decisions when it comes to sharing information with third parties.

The data is then shared for various purposes, all keeping in mind the best interests of the farmers. Some instances of data sharing include[86]:

Internal purposes

For JJ to better understand the impact of their work, provide agricultural advisories to farmers in their local languages and to plan for future strategies that can be implemented like climate resilient action plans.

External purposes

Sharing the data with researchers for drafting agricultural policies, to report back to corporate social responsibility funders for updates on projects, and general research for national development.

Image source: https://www.pexels.com/photo/brown-shovel-296230/
[85] Internal documents shared by Jaljeevika
[86] Internal documents shared by Jaljeevika

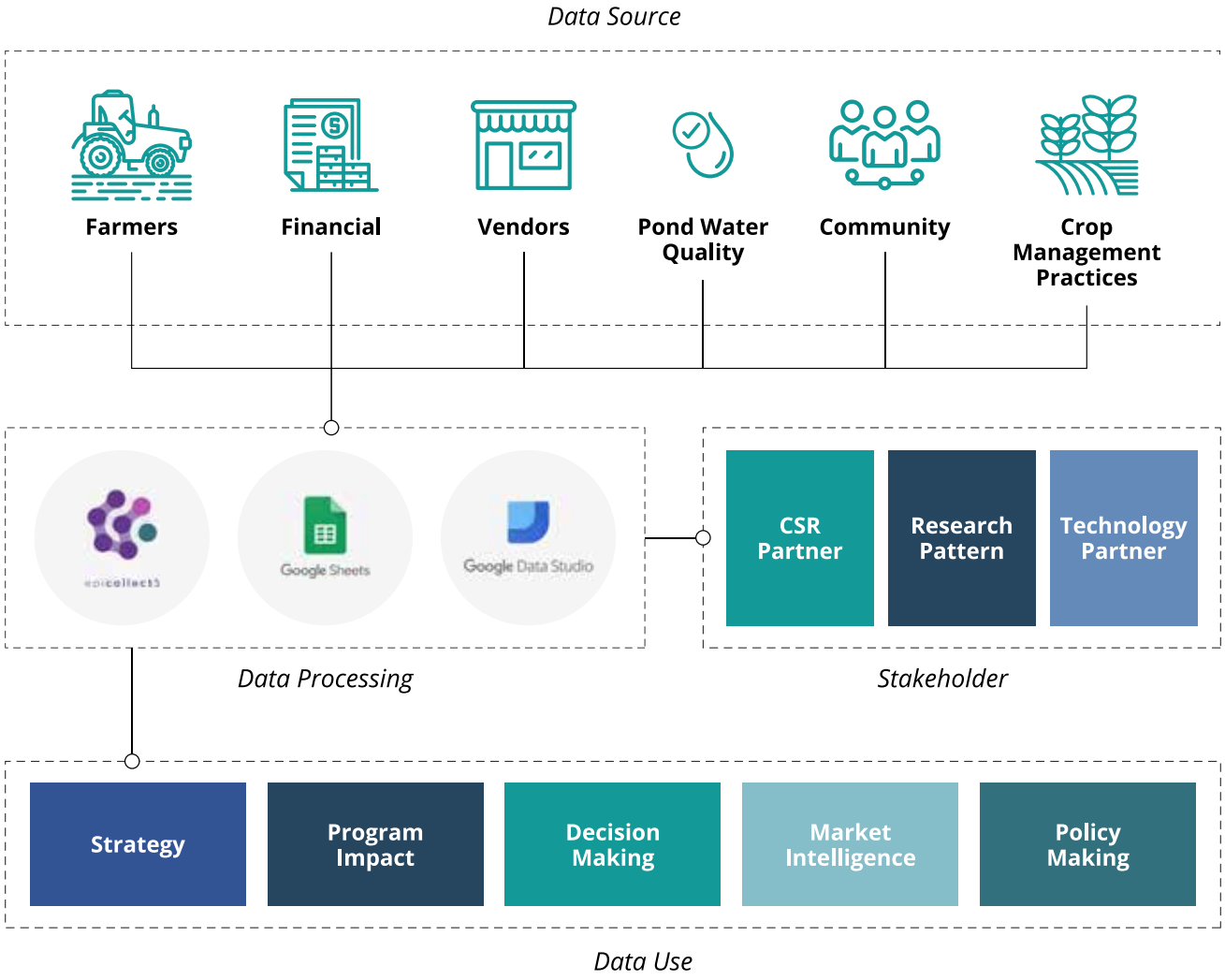


Image 3: Representing Jaljeevika's data collection and sharing process

Image source: https://www.freepik.com/premium-photo/asian-farmer-working-field-spraying-chemical_6831189.htm#page=3&query=farmer%20india&position=20

Methodology

For the purposes of this report, we used three methods to understand the working of JJ:

01. Interviews with internal members of JJ

We spoke to persons from the administrative department, who are in charge of the operations and management of the projects. With this, we were able to get an overview of JJ's plans to scale up their work.

02. Review of internal documents

This included their stakeholders from whom they are collecting data, the type of data being collected and shared, and the data flow planned for the IoT application. This gave us an overview of the underlying principles of their data collection application, the details of the technical architecture of their current data sharing systems and the policies adopted to protect the privacy of users.

03. Interviews with beneficiaries of JJ's work

We spoke to farmers, members of Self-Help Groups and farmer producer organisations, agents who are responsible for collecting information on behalf of JJ, and vendors. Through this, we were able to glean a holistic appraisal of JJ's work – from both independent users of their platforms and small agricultural businesses.

How will the Non-Personal Data Regulations (NPDR) affect the working of JJ?

JJ may be classified as a 'data trustee' under the regulations. 'Data trustees' are institutions that create, maintain and share community data, and work in the latter's interests. They can either be a government body, a not-for-profit organisation or can be formed by the coming together of members of the community. Data trustees have a duty of care over the data that they control and are obligated to set up grievance redressal mechanisms in case any harm occurs to the user's data. For example, harms that may occur in case of re-identification of an individual from the data shared. These trustees can identify opportunities to combine data from multiple businesses for the benefit of the community. The NPDR unequivocally declares all data trustees to be data businesses. They can charge a nominal fee for performing the function of data processing (and not data collection, as the NPDR assumes that it is a part of their business). This classification might lead to an increase of JJ's business costs and responsibilities because they will have to set up the infrastructure necessary to comply with NPDR. For example, since data is going to be mandatorily shared with the government, JJ must ensure complete confirmation of the farmer's consent during the initial data collection process.

Another possible complication will be that certain non-personal aggregated details of farmers, such as the crop yield from a state, might have to be mandatorily shared with the government, irrespective of the farmer's consent. This not only violates the right of the farmer to refuse or withdraw consent under the PDP Bill (though the NPDR allows data subjects to refuse consent, but at the cost of not getting their non-personal data anonymised,

which is a major privacy risk), but also makes it impossible for JJ as a data trustee to work 'in the interests of the community'. Both JJ and the farmers have little incentive to share data under the NPDR – the regulations talk about compensating trustees only for the data analysis process. It does not make any mention of the compensation to be given to communities as it looks at non-personal data purely from a market rather than a user-centric perspective. The increasing compliance costs incurred by JJ under the NPDR will hardly be balanced by this compensation. Besides, JJ as a trustee can ask for a 'nominal amount' for analysing this data. This might be tricky in a situation where there is a power imbalance between JJ and the data requester – if the latter has better negotiating powers, it might not pay the amount actually due to JJ for their work.

The difference in JJ's present and (possible) post-NPDR data sharing structure is that currently it has the autonomy to reject the sharing of data if the contract does not comply with its internal standards. Under the NPDR, if the data custodian or data requester is a government body, it may not be at liberty to exercise this power of the mandate to comply with the rules. If faced with the threat of penalty, JJ will have no option but to share the data with the requester.

JJ may also be classified as a data custodian under the NPDR, an entity that has a direct relationship with the data principal and collects, stores and manages this dataset. That JJ works directly with the farmer community makes the possibility of such a classification much higher. Custodians are obligated to share non-personal data if requests for this information abide by the standards set in the NPDR. This data must be shared in anonymised form, with due procedures undertaken to protect the user from any harm. Data custodians too have been declared to be data businesses, which means that if JJ is classified as a data custodian, it will have to mandatorily share data of farmers with the government. However, data custodians can only either be a government body or a private organisation. Technically, JJ is registered as an NGO and hence does not fall in either category. It remains to be seen how the NPDR will clarify such conflicts.

Image source: https://www.pexels.com/photo/person-reaching-for-banana-fruit-983466/

A critical analysis of JJ's role as a data steward

Irrespective of the type of classification, JJ's role will entail a stewardship angle. A data steward[87] is a trusted intermediary who works on behalf of users to manage their data without any vested interest. Considering this is an evolving model, there are many nuances to it that may be considered.

For one, a steward cannot always be expected to be a neutral body acting without any interests in the data. As a business model, this might not be profitable for them if they get a monetary compensation that is very little when compared to the services provided by them. JJ, whose work is primarily focused on improving the livelihoods of farmers by providing them with agrii-based advisory, might not be able to sustain on this model. There must be some form of monetary compensation that comes with this data sharing process which will be an incentive for them to continue their stewardship work. As an alternative, they can choose to run parallel businesses to make up for these costs. Future regulations on stewardship can set standards to ensure that the data stewards work for the community and their separate businesses do not conflict to the detriment of the farmers interests.

Considering its role considering technical nuances before agreeing to share data, JJ will either have to link itself to an external support system (for example, organizations providing legal advice on issues relating to privacy rights) or may have to undergo training to update itself on these issues. While JJ is aware of the existing regulations, it is still unsure about the implications it might have. This is where a pilot project would come to use - organizations like JJ will be able to better prepare themselves at the administrative and community levels.

Confirming consent before collecting data has always been important. With the possible introduction of the mandatory data sharing framework by the NPDR, data will have to be shared irrespective of consent for public good purposes. Therefore, getting the consent of the individual at the initial stages of collecting data is crucial. Currently, JJ, through its community agents, explains the terms and conditions of a contract to a farmer as a part of their consent process. But our conversations with the farmers showed that this might not always lead to an effective outcome. Due to their many years of work on ground, the farmers have developed a relationship of trust with JJ. This can lead to them agreeing to contracts merely based on this established relationship, rather than by actually taking the time to understand the details of the contract. This is a form of consent fatigue that can actually harm the community. JJ must streamline its consent process at the initial stages of data collection and emphasize its importance to the farmer.

Image source: https://www.pexels.com/photo/woman-picking-plant-on-field-916406/

[87] "Understanding data stewardship: taxonomy and use cases", Manohar S., Kapoor A. and Ramesh A., Aapti Institute, retrieved on September 19, 2021 from https://uploads.strikinglycdn.com/files/64aa4010-6c11-4d6f-8463-efaed964d7d9/Understanding%20Data%20Stewardship%20-%20Aapti%20Institute.pdf

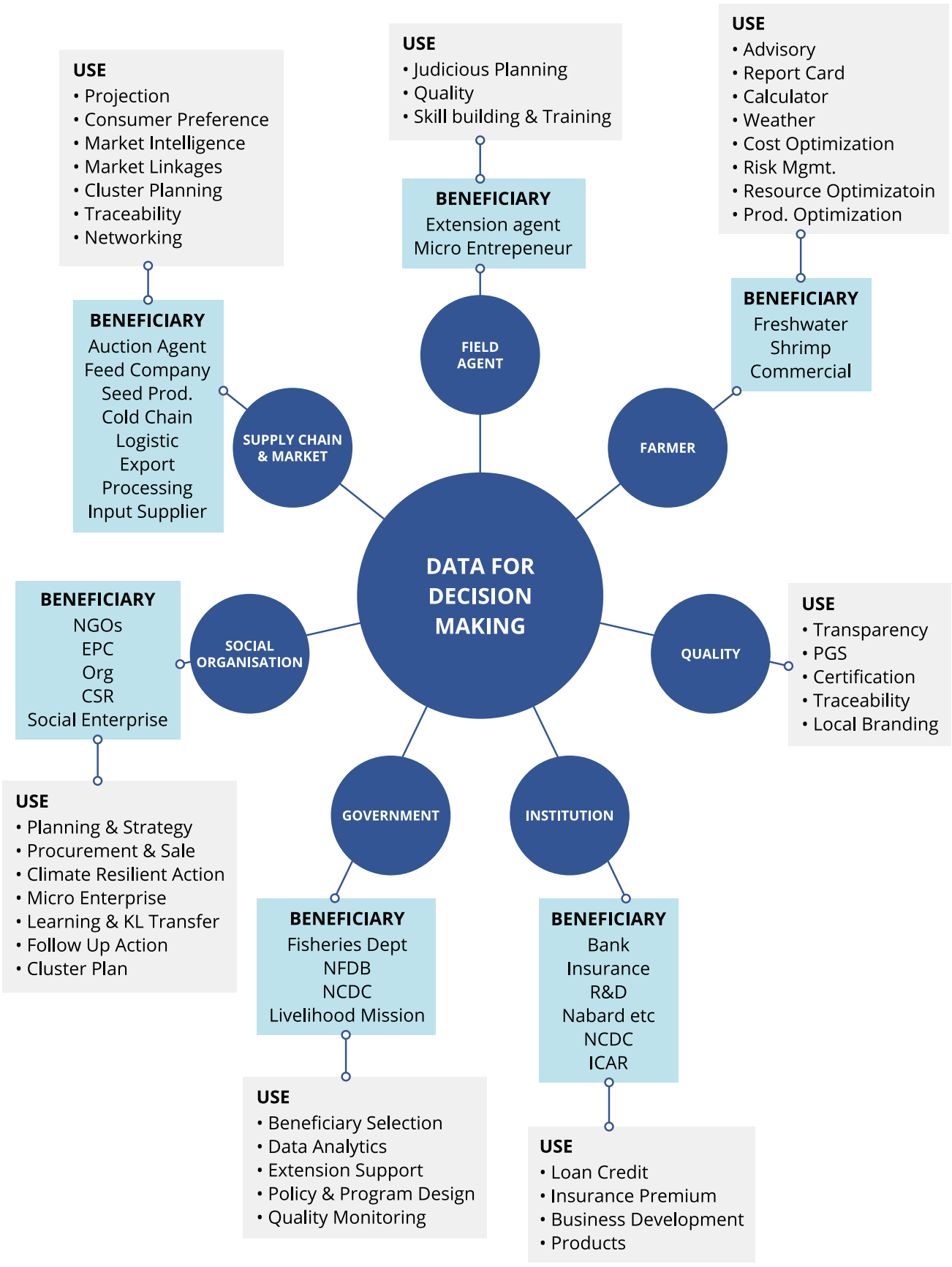


Image 4: Representing the current purposes for which Jaljeevika shares community data

According to its documents on data flow, JJ currently only shares data with three stakeholders - researchers, CSR partners and technology partners. With the NPDR, JJ will be faced with situations where it will have to negotiate with powerful parties like Government bodies and big businesses, for the sharing of data. In preparation for this, it is advisable for the organization to update its internal policies to include elements like non-negotiable standards for data sharing (example, the data requester must have data protection policies), and must incorporate principles of purpose limitation and data minimization (example, only the data necessary for the purposes of the contract must be shared). This can be in the form of data protection policies and can help improve its accountability to the community. The NPDR also says that data trustees can ask for a 'nominal amount' from the data requesters for the analysis process. A pre-established pricing policy would be of benefit for JJ.

Overall, JJ's current working structure already incorporates some of these recommendations - the changes might not be difficult to make, but are important as it undertakes its responsibilities under the NPDR. They are also making commendable efforts to update themselves on data protection and right to privacy policies, giving them ample amount of time to prepare for what the NPDR may bring.

Image source: https://www.freepik.com/premium-photo/indian-farmer-cotton-field_9665978.htm#page=1&query=farmer%20india&position=15

Annexure B

List of foreign policy documents referred to in the report

The United States' "Privacy and Security Principles for Farm Data" (Principles for Farm Data)

The European Union Code of Conduct on Agricultural Data Sharing by Contractual Agreement (EU Code of Conduct)

The New Zealand Farm Data Code of Practice, Standards and Accreditation (NZ Data Farm Code)

The Australian Farm Data Code

General Data Protection Regulation (GDPR)

Annexure C

List of interviewees

NAME OF THE INTERVIEWEE	INSTITUTIONAL/ORGANIZATIONAL AFFILIATION
Adam Kriesberg	Assistant Professor at Simmons University's School of Library and Information Science
Can Atik	PhD scholar from Tilburg University
Leanne Wiseman	Associate Director of Australian Centre for Intellectual Property in Agriculture (ACIPA)
Nachiket Udupa	Mazdoor Kisan Shakti Sangathan
Ranjini Basu	Agricultural Policy Officer, Focus on Global South
Tanay Mahindru	YLT Fellow
Yogesh Patil	CEO of Skymet Weather Services Private Limited

aapti institute 2021

Aapti is a public research institute that works at the intersection of technology and society. It examines the ways in which people interact and negotiate with technology both offline and online.

contact@aapti.in

www.aapti.in | www.thedataeconomylab.com