Input to EU's draft Data Governance Act

IT for Change[1]

IT for Change (www.ITforChange.net) is an India based NGO, with global footprint, and a Special Consultative Status with the UN-ECOSOC. We have been represented in many committees and working groups at the global level, including in the UN, as well as in India – the latest being the Government of India's Committee of Experts on Non-Personal Data Governance Framework.

The draft Data Governance Act (DGA) is a welcome move by the EU to enter into the area of economic governance of data, beyond human rights based governance embodied in the GDPR. It is positioned as the first step among a series that are envisaged in the EU's Data Strategy. On the eve of the release of the draft DGA, the European Commissioner for Internal Market, Thierry Breton, described Data Strategy's aim as making EU "the most data-empowered continent in the world" and tip the scales away from Big Tech.[2]

The Data Strategy itself puts the intention more moderately; to ensure that by 2030 EU's share in the data economy rises to at least the level of its (otherwise) economic weight. This itself is no small task in a digital economy where, as per UNCTAD's estimates, US and China account for 90 percent of capitalisation of top 70 digital platforms globally.[3] This bipolar global concentration of digital and data power seems set to further intensify.

The Data Strategy asserts that "currently, a small number of Big Tech firms hold a large part of the world's data". These are of course all US and Chinese firms. Discussing different data models, it observes: "In the US, the organisation of the data space is left to the private sector, with considerable concentration effects." "China has a combination of government surveillance with a strong control of Big Tech companies over massive amounts of data without sufficient safeguards for individuals." In order to release Europe's potential, it is necessary to find a third, "European way, balancing the flow and wide use of data, while preserving high privacy, security, safety and ethical standards".

The draft Data Governance Act is expected to "facilitate data sharing across the EU and between sectors to create wealth for society, increase control and trust of both citizens and companies regarding their data, and offer an alternative European model to data handling practice of major tech platforms."

A key objective of the draft Data Governance Act, quoting it, is the "establishment of the common European data spaces in the near future, in line with the EU recovery plan". It proposes to establish "domain-specific common European data spaces, as the concrete arrangements in which data sharing and data pooling can happen"."Common European data spaces will ensure that more data becomes available for use in the economy and society, while keeping companies and individuals

---

1 For any further information or clarification, please contact parminder.js@itforchange.net .

2 https://techcrunch.com/2020/11/24/europes-data-strategy-aims-to-tip-the-scales-away-from-big-tech/

3 UNCTAD's Digital Economy Report, 2019. https://unctad.org/news/global-efforts-needed-spread-digital-economy-benefits-un-report-says

who generate the data in control."[4]

We will judge the draft Data Governance Act against these key objectives. We can arrange them in three parts: (1) Sufficiently well provided sectoral data spaces do come into existence, that can provide the basis for the European digital industry to stand up and compete with the incumbent US and China giants. (2) These data spaces are governed effectively, for efficiency and fairness. (3) Control over data is kept or shifted to be with those who actually generate data.

On objective (1): the draft DGA only provides a framework for voluntary sharing of data. Having established that most of world's data is with a few global Big Tech firms, it is not made clear why this biggest part of the required data will get shared voluntarily, when Big Tech firms consider exclusive access to and use of such data as one of their main business advantages – if not 'the' main one. Without access to society's data hoards held by a few global Big Tech firms, there can be no effective sectoral data spaces that can adequately feed European digital and AI businesses. Draft DGA seems to be framed with the supposition that effective governance of data spaces itself will be enough to attract all the needed data into sectoral data spaces. This would not happen, especially with regard to the most important and extensive data that is with the Big Tech. In fact, the provisions of DGA may help Big Tech more easily access society's data that it is currently not collecting.

On objective (2); even if we take the draft DGA as only a regulatory measure, and deliberately not establishing any data access rights (which the draft DGA says will be done by a Data Act in 2021), it is strange that the draft DGA almost completely circumvents actual existing, mainstream processes of data collection, use and sharing, putting them beyond its pale. What the draft DGA's provisions aim to regulate are some ideal-types of neutral data exchanges. These hardly exist currently in any appreciable measure. Regulators seem to expect such ideal types to become mainstream just because a good governance framework now gets provided for it. We consider such hopes also to be very unrealistic.

We are of the opinion that EUs half-hearted approach of 'build it and they will come' is simply not up to the challenge of making EU "the most data-empowered continent in the world" and tip the data power balance away from large tech companies.

The objective (3), of putting control in the hands of those who generate data, is also not met, because almost all provisions of draft DGA concern voluntary practices that main data actors are under no obligation to observe. Some 'possibilities' of controlling individual data, by natural and legal persons, do get provided, but without making them mandatory for all those who deal with data. The responsibility remains upon individuals and small businesses to go out and subscribe to new data sharing services instead of dealing directly with Big Tech. That there may not be enough incentive for them to make such a shift, especially given the huge repository of carrots and sticks at Big Tech's disposal, is a well-known digital economy fact that is conveniently avoided.[5] Meanwhile, there is nothing on how one of the most important and valuable categories of data,

[4] https://ec.europa.eu/digital-single-market/en/european-strategy-data#:~:text=Common%20European%20data%20spaces%20will,and%20societal%20progress%20in%20general.

[5] This is like keeping GDPR compliance optional and expecting that when the market will have both GDPR compliant and certified players and those that are not, people will automatically flock to the former. Draft DGA perhaps expects data monetisation to come into the picture to make data sharing services the preferred option. Data monetisation, especially as against competing benefits that can be provided by Big Tech for dealing directly with them, is a complex subject. But what is generally known about it does not give hope that draft DGA's expectations will hold.

collective and aggregate data, including of anonymised kind, can be controlled by those who together generate it. One promising idea of 'Data Cooperatives' is presented, but the relevant provision is half-hearted, and largely empty with regard to real possibilities of its application.

The rest of the paper discusses these issues at greater length.

Voluntary data sharing leaves out the biggest data hoards

If the objective, as stated by the European Commissioner for Internal Market, is indeed to become the leading data continent and tip the balance of data power away from (US/China based) Big Tech, it is not clear how this can be achieved without in some way mandating that Big Tech shares the data that it collects from and about EU's society.

Developing extensive and efficient voluntary, including price-based, data sharing frameworks and arrangements may, in fact, just allow Big Tech to access data that it is otherwise not able to lay its hands on. This further enhances its data power. Coming from India, if we may employ a colonial age analogy; the infrastructure of railways and ports created by the British Empire got employed to denude India of its resources and further its de-industrialisation. Simply creating systems or infrastructures for data sharing therefore may not cause digital industrialisation in the EU, it can even have the reverse effect.[6] Big Tech has the money, technical expertise, and persuasive means to become the biggest beneficiaries of voluntary data sharing arrangements, without contributing anything, or, at any rate, anything substantial, in return. EU needs to give attention to this very likely outcome of its current data policy orientation.

We understand that the EU's current approach is based on the premise that the forthcoming second phase of dataficiation and digitalisation will be different from the first phase in a few important respects, that favour the EU.[7] Instead of coming from consumers, more data is likely to be industrial and with businesses. Second, more data processing would increasingly take place at the edges and not centrally.[8] An industrial/business data generator is relatively more capable of exercising data related choices, which, with close-to-edge data processing, become easier to undertake. Building on these premises, the draft DGA mostly attempts to provide a framework for an effective exercise of such choices.

There may be some truth in what a new or second phase of dataficiation and digitalisation may signify. But EU policy makers seem to be taking this argument too far, and trying to solely, or mostly, rely on it. This is perhaps because it allows them to conveniently avoid the most important imperative of finding ways to access the main hoards of society's data which is with a few global Big Tech firms. An over-reliance on this one proposition of the second wave of dataficiation being different, and inherently EU favouring, is based on an inadequate understanding of platformisation. As platformisation proceeds to physical sectors, involving existing industrial and business users, it structurally transforms these sectors. Platforms taking up key data controls is a key element of such a transformation. It then does not remain that important whether data is processed closer to the edges or centrally; what matters is who controls such data. (We understand that the proposed Data Act will give some attention to this issue, to which we will come soon.) Larger industrial players may theoretically be better able to retain and exercise data controls, but this may not automatically

[6] We do acknowledge here that EU is making efforts in other areas of digital economy as well, like cloud computing, AI etc., but the essential argument of Big Tech taking disproportionate benefit from data sharing systems still hold.

[7] cite

[8] These premises are stated in the draft DGA, and often articulated by top EU leaders.

be the case. The way European car manufactures are clustering around US (Waymo) or Chinese (Baidu's Apollo) digital mobility platforms is a cautionary tale that should be heeded, and provided for.

The draft DGA as well as the Data Strategy say that issues of rights about data access and use will be dealt by a subsequent legislation, Data Act, expected in 2021. It will be eagerly awaited. However, it seems that its focus will mostly be on industrial data that is 'co-generated' (we understand, between the industrial machine owner and digitalisation/datafication service provider). Industrial data has become a somewhat inward looking and myopic obsession of EU policy makers, which misses the points of (1) how platformisation transforms any sector that it touches (see the above para), and (2) that industrial data though important is just one segment of society's data, there are services related data and other general non personal data throughout the society, as everything gets digitalised.

Having recognised that most of society's data is with (non EU) Big Tech, it is discordant to see the Data Strategy say that even the Data Act will not address the data power of Big Tech. That part is supposed to be left to a separate process specifically addressing the "high degree of market power of certain platforms". It is unclear, but perhaps that process is what has resulted in the draft Digital Markets Act, also currently open for public comments. We will separately respond to the draft Digital Markets Act (DMA). Meanwhile, the draft DMA does provide some very important implicit rights to business users of large platforms of accessing and using 'their data' with these platforms. These rights, expressed as mandated prohibitions against data abuse as well as data access provisions, can have far reaching implications, as we discuss a little later. Suffice to point out here that the draft DGA prima facie remains oblivious to such new rights of use and access, and how these may require enabling frameworks to be effectively operationalised.

Any general data governance framework, as sought to be laid out by the draft DGA, must also be mindful of mandated data access and sharing possibilities, and suitably enable them. Some such possibilities are provided by the draft DMA and more are likely to figure in the Data Act. Frameworks that give data access rights, and mandate data sharing, should precede and be subsumed in a comprehensive framework of data governance. In defining 'data sharing' under it as only 'voluntary data sharing' the draft DGA structurally keeps out all mandated data sharing from its purview. Any data intermediary service, on the ground, may need to simultaneously mediate voluntary as well as mandated sharing of data of various kinds. The draft report of India's Committee on Non-Personal Data Governance Framework recommends data trustees that will mediate mandatorily shared data, likely, along with voluntarily shared data.[9] It is surprising that the draft DGA leaves no room for such a possibility – now or in the future, when the EU may have some mandatory data sharing provisions. It defines 'data sharing' to only mean voluntary data sharing, and requires that "the data sharing service should be provided through a legal entity that is separate", meaning that the same entity cannot mediate mandatory data sharing. The draft DGA is evidently completely blind to mandatory data sharing even as a future possibility.

This exhibits why a general framework for governing data spaces should follow rather than precede frameworks of data access rights that ensure that such data spaces actually get populated by all the required data. Getting this sequence wrong evidences the tendency of EU policy makers to keep circling around what is one of the biggest and most urgent contemporary economic governance problem without undertaking upfront what is strongly indicated from their numerous deep analyses as well as political statements. Such a failure of political will to get on with the necessary data

[9] https://static.mygov.in/rest/s3fs-public/mygov_160922880751553221.pdf

policies and frameworks will not stand the EU in good stead. Time may be fast running out for making good on EU's ambition to be world's top data and digital continent, or perhaps even to substantially survive in the global digital race.

Furthermore, not addressing the central question of economic rights to data first but proceeding with lame-duck frameworks for economic governance of data has this very dangerous result of entrenching default models of data ownership and rights that strongly favour the biggest data collectors and hoarders, i.e. the Big Tech. An example of this is how the draft DGA introduces a new term 'data holder', the actor who is central to the provisions of the Act. In its objectives it seeks to empower those 'who generate data' but in its substantive provisions all rights and claims are with the category not of 'data generators' but of 'data holders'. Use of this term strengthens the de facto data-related rights and ownership of data collectors and hoarders,[10] by implicitly recognising them in a legal instrument of economic governance of data. For personal data, some rights at least always remain with the data subject, wherever that data may be at any given moment of time. But no clear rights whatsoever exist around non personal data, including anonymised personal data. With no other legal encumbrances, as 'holder' of such data, big data platforms are able to exercise complete legal rights over all such non personal data under the draft DGA. This serves to give implicit legal cover to what so far is only de facto ownership of most of society's non personal data. Data holders "allow(ing) the use of *their* non-personal data" (emphasis in italics added), is an instance of the kind of problematic formulations in the draft DGA that we refer to here.

This way too, the draft DGA may be contributing to the precise opposite of what it set out to do – tame the data power of Big Tech, promoting EU's digital and data industry in relation. Explicit frameworks for rights of access and use of data are therefore urgently needed in favour of smaller, less-powerful, players in a digital economy. These rights may take the form of flow of relevant data into common data spaces, that can provide the required safe and appropriate data access. Substantive data rights frameworks, mandating access to relevant data (what to share), must precede a general framework on process oriented data governance (how to share) as provided by the draft DGA.

Governing data spaces – Shirking reality while aiming at marginal ideal-types

In its detail, sophistication and nuance, the draft Data Governance Act is an exemplary legislation, as EU's efforts of this kind generally tend to be. Its treatment of access to public sector data, which is not of 'open data' nature, would enable important and substantial contribution of data to EU's proposed sectoral data spaces. Ensuring trustworthy mechanisms for voluntarily proffering data for specific public good purposes, or data altruism, is also commendably achieved in the draft DGA.

These however are obviously the easier things to do. The real, and admittedly difficult, issue is to suitably address commercial practices around data. But, even in the voluntary transactions space, the draft DGA meticulously avoids regulating mainstream commercial data practices.

---

[10] The Progress Report of Expert Group for the Observatory on the Online Platform Economy, Work stream on Data, observes that "At present, control over data in the data economy is largely dominated by de facto technical control and contractual rights. In the absence of any specific legal provisions, this leads to a de facto ownership, meaning that businesses or organizations in the position to collect data are able to exercise more or less exclusive control over it (albeit within the legal frameworks set by the GDPR and other relevant legislation)." https://platformobservatory.eu/app/uploads/2020/07/ProgressReport_Workstream_on_Data_2020.pdf

After referring to lofty ideals of "alternative to the current business model for integrated tech platforms" through a 'novel European way of data governance', that would inter alia ensure a "separation in the data economy between data provision, intermediation and use", the draft DGA comes up with a damp squib. It offers its 'good governance' 'if' data actors chose to employ certain ideal-type data practices. This ideal-type consists in making "a direct relationship between data holders and data users", without any kind of additional enhancement services, and excludes cloud computing service providers. It is the cloud computing service providers like Amazon, Microsoft and Google that run the most popular data exchanges today. Most data intermediation arrangements tend to provide some additional services, like, refinement, aggregation, analytics, etc.

To clarify again, the new regulation does not push data sharing and transfer practices towards any ideal-type or 'good practices'. It just offers its governance framework 'if' the ideal-type is adopted. And this 'ideal-type' is cut very narrowly to overwhelmingly exclude most current data practices. To avoid coming under the provisions of the draft DGA, one just has to avoid being this narrowly tailored ideal-type – which seems easy to do, if one is already not so.

This leaves the draft DGA to only cover some niche data intermediation services,[11] bypassing the mainstream data practices. Apparently, it is hoped that services that voluntarily accept the somewhat stringent conditions of draft DGA will be seen to have quality, neutrality and trustworthiness. This would attract users and customers and thus these services will become popular, even mainstream. Good data practices are supposed to become a competitive advantage, and the draft DGA offers help in this process by quality marking. This carries a strong whiff of how GDPR was earlier supposed to have an industrial policy angle, whereby good data practices of privacy protecting EU services were expected to become their competitive advantage against foreign Big Tech. We know that this really has not happened; in fact the GDPR probably worked to Big Tech's advantage vis a vis European firms.[12]

This shows that anything proposed to have an industrial policy interest – as the Data Strategy and its various elements obviously have – should better have its objectives and game-plan clearly cut out. This is especially so in the high stakes digital and data contest. Expecting optional good governance practices by themselves to serve the industrial policy objective of advantaging the complying domestic firms is mostly unrealistic.

It remains unclear how such provisions will contribute to breaking the Big Tech's model of integrated data practices, an express objective of the draft DGA. The ample repository of carrots and sticks at Big Tech's disposal to retain and further entrench such a model is too well known to have missed the policy makers' notice. And, how the desired structural "separation in the data economy between data provision, intermediation and use" is achieved just by prohibiting any general data intermediary from providing data value add services but not doing the reverse; banning data value add service providers from undertaking data sharing and/or directly sourcing data from data generators?

The integrated data model of Big Tech will be broken, and actual structural separation across the data value chain ensured, if Big Tech, as provider of data-based services, is not only (1) prohibited

[11] Reference to MyData service comes in the Data Strategy, and often in literature on data sharing. It is not clear that the reason that MyData, and similar services, have not risen quickly and displaced mainstream data practices has to do with their trust deficit and lack of good governance, which kind of perceived gap alone the draft DGA aims at filling.

[12] https://www.wsj.com/articles/gdpr-has-been-a-boon-for-google-and-facebook-11560789219

from sharing data outside its own uses, but also (2) mandated to collect data only through third party, DGA compliant, independent data intermediaries, and never directly from data generators/ data subjects. This is what would actually result in what the draft DGA says it seeks to achieve – an "alternative to the current business model for integrated tech platforms". If such would be too far-reaching a regulation disrupting existing practices too extensively in one go – an understandable consideration, this issue should at least have been presented and discussed, and an appropriately gradual approach taken. Our organisation, IT for Change, did a paper last year, 'Breaking up Big Tech: Separation of its Data, Cloud and Intelligence Layers',[13] which discusses such mandatory structural separation across the key elements of data value chain. It is time EU begins to think on these lines, and introduce gradual, but firm, measures in such directions.

A primary purpose of the draft DGA is to "put those who generate the data in the driving seat — moving away from the current practices of the big tech platforms". The draft DGA addresses this issue with respect to two kinds of data generators, individuals and business users. What is missing here is a group or community as data generator. The greatest value of data can in its social relational form.[14] Data is very often generated on platforms by a group activity and not individual. It may be data about a pattern of behaviour across similar actors, or it could be about interactions among them. There is increasing literature on collective privacy.[15] But there could also equally be collective positive rights, concerning collective benefit from a group's aggregate data. Just because the latter cannot be identified with any one individual is not an adequate basis for the economic value of such data to entirely rests with the data collector or holder – as treated in dominant data practices. This is one of the biggest blind spot of EU's data rights and data governance approach. Data sharing services under the draft DGA are expressly not supposed to undertake data aggregation, and then deal with such aggregate data. A group therefore cannot get its data aggregated and anonymised by such a third party neutral service and exercise control over such data – for preventing harm, and gaining benefits. Dealing with the all valuable aggregate and anonymised data is thereby – even in letter and intent – completely left to mainstream data practices under control of Big Tech.

To take an example; there could be much data about farmers and farming in a region, that can be collective at source – drone images of general crop and soil patterns in a region, or may arise on individual farms but best employed in aggregate forms to find patterns and intelligence. Some efforts are afoot to give farmers individual control over data originating from their farms and their activities, like in the 'EU Code of Conduct on Agriculture Data Sharing'.[16] But it is never clear how such common and aggregate data will be controlled by data originators[17] or data generators – which control has obviously to be collective, and is increasingly all important in the data vale game. The draft DGA unfortunately provides no substantive help in this matter.[18]

The empty provision of 'Data Cooperatives'

Collective data and its collective management does find an implicit mention in the 'Data Cooperatives' part of the draft DGA, in a passing and tentative fashion. There is a reference in the

[13] https://datagovernance.org/report/breaking-up-big-tech-separation-of-its-data-cloud-and-intelligence-layers
[14] See for instance this paper at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3727562
[15] https://www.springer.com/gp/book/9783319466064
[16] http://www.fao.org/family-farming/detail/en/c/1370911/
[17] A term employed in the 'EU Code of Conduct on Agriculture Data Sharing'.

[18] Not just the data sharing services related provisions, but also the public sector data sharing provisions of the draft DGA are blind to any – existing or future – new data rights, only recognising GDPR related rights and traditional IP rights. Encumbrances put by collective rights against harm may be relevant in this case.

recitals to ‘data pertain(ing) to several data subjects within … (a) group’. This could be a group of individuals, or, evidently, also small businesses. There is also a suggestion regarding mechanisms for group members to take collective decisions about such group data. However, it is then made clear that there are only individual rights in any data involved (as per the GDPR) and no collective rights subsist. This is quite anomalous. Individual rights can only be exercised to the extent data remains personally identifiable; but what happens to data rights beyond that line? If ‘data pertaining to several subjects within a group’, or aggregate data, is anonymised, as it often is for safety and other reasons, how can individuals give individual consent or permissions for such data? There can only be collective control and collective consent or permission. Such collective data controls could be rights-based – which being stronger are much preferred, or just contract-based – which would fulfil limited purposes and do not constitute effective controls. The draft DGA avoids such important matters.

[Without getting into hair-splitting, the confusion and half-heartedness with regard to aggregate and collective data is evident from the following. Recital 24 on ‘Data Cooperatives’ speaks of “solving disputes between members of a group on how data can be used when such data pertain to several data subjects within that group” which suggests collective decision making related to aggregate data – which is the only way such data should be controlled. But the substantive section 1 (c) of Article 9 about ‘Data Cooperatives’ only mentions individuals or small businesses “confer(ing) the power to the cooperative to negotiate terms and conditions for data processing before they (apparently, individually) consent”. There is nothing about any collective consent, permission or control, as seems implied in the language of the recitals. How can persons or businesses individually consent about data already aggregated and anonymised – is it a matter of veto, whereby if even one does not consent, collective data cannot be shared or used? Cooperatives normally do not run with such extreme ‘a single veto stops everything’ provisions. But there does not seem to be the will and/or conceptual boldness and innovation to deal with such real data governance issues. It would have been useful if knowledge and practices from real world cooperatives were brought into play here. This is why we say that this concept of ‘Data Cooperatives’ is introduced half-heartedly and serves not much real purpose.]

Insertion in the draft DGA of the concept of ‘Data Cooperative’ seems largely decorative. There is nothing that provides new support or enablement to a Data Cooperative. Rather, new notification-related and other strong conditions have been put on any such activity – which may simply be a burden. Altogether, this provision is neither here nor there, and appears mostly as an empty slogan. Same actors that may form a Data Cooperative under the draft DGA could anyway have associated in a group and decided together on how they will allow their data to be used, if it was to be based only on existing individual rights, under the GDPR, or just on the limited remit of private contracts. Under the draft DGA, any such activity has to now meet many new standards and qualifications. It is not clear how this facilitates rather than burdens data generators seeking to act cooperatively to exercise data choices.

It would have been different if, apart from helping coordinate individual consents – which groups can in any case do, some rights were provided vis a vis collective data (the kind which is mentioned as ‘data pertaining to many data subjects’). Such collective data requires rights based protection for the same reason that individual data does, because (1) the parties concerned are much weaker as against the data collecting and using parties, and (2) certain claims, like related to some kinds of collective harms, have to be non-negotiable. Private contracts based aggregate data offerings and transactions do not meet these requirements.

Let us explain the need for collective data rights with a few examples. Consider a hypothetical community of patients with a rare disease. Each member of this community can of course determine respective choices about personal data. Members can also coordinate over their existing community channels how best to do so. They may be with a common personal data sharing service provider, or different ones – if any at all. In either case, the final choices as made – through consultation or not – can be communicated and exercised through such a service, or directly. It is not clear what does the 'Data Cooperative' provision of the draft DGA add here. Is it that it is now illegal to so coordinate their individual choices through existing community channels without providing a notification to the authorities, and meeting other conditions, as required in the draft DGA? That would be an unnecessary impediment to a very genuine activity, especially when the draft DGA brings no new benefits to the community.

What has to almost be equally of concern is how the community of patients deals with data that is anonymised across its members. Such data could be very useful in developing treatments for their common disease, or may even have other uses beyond that patient community? Also, they will like to prevent any harm caused to them, or elsewhere, through use of their non personal data. Currently, this group has no collective rights over their collective data. Once anonymised, it can get completely out of their control, and be used in any way that the data collector or holder deems fit. In fact, as per the draft DGA too all rights to such data seem to be with the 'data holder' with no encumbrances.

Even if the patients' community were to employ a collective approach to share their anonymised data with contractual conditions about its further use– using a Data Cooperative or otherwise, it does not sufficiently protect it. This is for the same reasons that individuals need explicit data rights beyond just consent based contractual protection. Such collectives also consist of parties that are much weaker, individually and even collectively, against data collecting commercial parties in transacting their data. Further, many possible harms and benefits may not be able to be pre-considered or otherwise possible to enter precisely in a contract. As with individuals, a rights based 'duty of care' and 'entitlement to benefits' with respect to their data should subsist as a collective right for any data group or community – which right can then be operationalised in various ways. It is after all their data, and provides valuable information and intelligence about them.

The 'Data Cooperative' provision provides no such additional enablement for protection of collective data interests, and just puts new regulatory burden on any activity that attempts coordinated data consents or permissions. We could therefore not understand why this provision is at all there. If intended to protect members from corrupt or exploitative management of cooperative bodies dealing with data, this may be rather premature because there are hardly any data cooperatives in existence yet. Effort at this stage should focus on bringing them into existence, and enabling them with collective rights and other collective possibilities with regard to individual and collective data.

Data Cooperatives employing collective rights provided by Digital Markets Act – A test case

Despite the draft DGA's apparent blindness to such possibilities, can its 'Data Cooperatives' provision still help operationalise collective data rights that may be provided in the future, or exist elsewhere? Interestingly, one such right implicitly figures in the draft Digital Markets Act (DMA). Business users of very large platforms (notified as gatekeepers) have been provided rights to their individual as well as aggregate data, in terms of prevention of harm (such data cannot be used in competition) and of access and use (presumably to enhance their existing activities, but more

creative collective uses are possible, as will be discussed shortly). Since aggregate data can mostly not be denominated individually, the implied right to it must be collective, involving all business users whose data it is together. This group would normally consist of all business users in a similar relationship with a particular platform.

Such similarly placed business users can employ a 'Data Cooperative' – though not clear why they need the DGA for this, apart from the additional burdens it now puts on any such effort – to collectively bargain and ensure that their collective data is not used in a manner that involves competition to them (the one specific harm that the draft DMA protects against). Interestingly, the platform using such data to develop any kind of automation, which has the effect of reducing or eliminating the role of such business users, would evidently fall foul of this provision. Meaning, Uber, for instance, cannot use data generated in the course of its cab drivers providing services using the platform to either itself develop an autonomous automobile service, or provide such data to others who may be planning to do so. This is a far reaching prohibition given the current plans and activities of digital automobile service providers.

The new right to access and use their individual and collective aggregate data, provided by the draft DMA, can be even better and more innovatively employed by business users of platforms. This can be done through forming a 'Data Cooperative' – but again, there is no special enablement provided by the draft DGA in this regard. Collective and individual data generated from the activities of business users and their end customers is the main business asset of a platform. Business users having such data under their collective control can completely transform their relationship with the respective big platform. Business users' collective bargaining power shifts so substantively, that they may practically be able to seek co-determination[19] of major business decisions taken by the platform. Taking Uber's example again, if Uber drivers have collective access to such data and could use it collectively through a 'Data Cooperative', they can practically come to co-own the Uber platform[20] – on the pain of running a possible parallel business or palming off data to Uber's competitors.

[One promising element common to the draft DGA and draft DMA is the growing, but still less than clearly articulated, recognition of the status and rights of those who generate data – directly, or through their activities – as the most important actors in the data value chain. This conception if taken forward appropriately, and strengthened with explicit economic rights for such actors, can help shift data power away from where it resides currently, almost all with data collectors – which is mostly the Big Tech. Much more work, however, needs to be done in this regard, including making explicit the implicit ideas and conceptions, and shaping laws and regulation on that basis. It also needs to be established whether it is the 'act of data generation' or 'being the subject of data' – with the concept of subjecthood extended beyond individuals to legal persons and groups/communities – that provides the most appropriate basis for holding data rights.[21]]

---

19 A concept in practice with legislative backing in Germany. See https://en.wikipedia.org/wiki/Codetermination_in_Germany#:~:text=Codetermination%20in%20Germany%20is%20a,in%20the%20Mitbestimmungsgesetz%20of%201976.

20 Such possibilities have been discussed in this 2020 paper on 'Economic rights in a data-based society' https://publicservices.international/resources/publications/economic-rights-in-a-data-based-society?id=10819&lang=en

21 Our view is that instead of 'data generator' – which implies some kind of an active role that often is not the case for much data – the key data actor should be seen in terms of being the 'data subject' – that whom data is about. This provides a better basis for entitlement to control over such data as it reveals intelligence about the subject. Such data subject can be an individual, business entity or a collective of either.

If adequately supported in law, structures like 'Data Cooperatives' can go a long way to enable control of data as an individual and collective asset of those whose activities generate it. Such control can have transformative possibilities, as discussed above. However, instead of providing enabling provisions, like collective rights, injunction against refusal to deal with Data Cooperatives, etc. – like trade unions have statutory rights based enabling legal support – all that the draft DGA does is to put any Data Cooperative under substantial new obligations. It is not clear how this furthers collective efforts of small actors to get a fair and beneficial treatment with respect to their data. The draft DGA seems to show no awareness of the collective action/rights provisions of the draft DMA, or any possible ones that may arise in the future, from the Data Act or elsewhere.

Meanwhile, it may be stated that the draft DMA provides, only implicit, collective data rights for a very limited set of actors – business users of notified gatekeeper platforms. Many platforms and other digital firms fall outside the definition of such gatekeeper platforms, against whom too small businesses and commercially acting individuals require individual and collective data protections and enablement. If an Uber like giant platform should be prevented from unfairly using drivers' data, so should also be a city based local transport platform. Further, even more significantly, it is discordant that while the draft DMA institutionalises collective rights to their aggregate data for business users of platforms, no such collective rights exist anywhere for end users of platforms, or otherwise for groups and communities of natural persons. All this makes the case for a wider approach to substantive collective data rights.

We discussed earlier the case of a community of patients with a rare disease. Farmers in an area may similarly want their aggregate and other non personal data to be used in a rights based manner only for purposes that they would like to define. There is no such legal provision or right anywhere currently. Residents of a locality may similarly want collective rights to their non personal data, including anonymised data – right against harm, as well positive right to best beneficial use, and benefit sharing. A properly institutionalised rights-based form of 'Data Cooperative' will be immensely useful in all such cases.

End note

The above discussion inter alia shows why process-oreinted provisions of data governance should follow, or come along with, laws that provide substantial economic rights around data. And there should be mutual recognition and harmony between these two kinds of data governance rules and laws. By putting the cart before the horse, the draft DGA achieves little if anything as far as mainstream commercial practices around data are concerned. (Issues of public and altruistic data sharing being relatively smaller issues, and certainly rather more straight-forward and easier to address. These important parts of the draft DGA are nevertheless much appreciated.) One understands that the draft DGA follows the design laid out by the Data Strategy, and therefore there may exist some level of internal understanding that ties things up. To that extent – pending what is unveiled by the Data Act – the Data Strategy itself may have strong deficiencies in considering actual legal means required to achieve what it describes as its key objectives.

It may be mentioned here that the earlier referred draft report of the Indian government's Committee on Non-Personal Data Governance Framework[22] provides extensive recommendations with respect to 'a community's right to its data'. It details how such a right can be operationalised – both to prevent collective harm, and to maximise benefits to the community concerned.

[22] https://static.mygov.in/rest/s3fs-public/mygov_160922880751553221.pdf

Most importantly, it seeks to ensure that the largest store of society's data with Big Tech actually feeds into a country's much required sectoral 'data spaces' or 'data infrastructures'. It also has explicit provisions on how groups and communities can protect themselves against collective harm. The draft report further requires all data related businesses to provide transparency about their data practices, and not just a select few. These recommendations therefore plug three major shortcomings of the EU's data governance approach.

The Indian document however does not address the individual and collective data rights of business users of platforms that are so well dealt by EU's draft Digital Markets Act. It also does not provide specific means for safely, easily and fairly accessing public sector data, and altruistically provided data, for society's overall needs, something achieved well in EU's Draft Data Governance Act. The Indian document also remains deficient in providing detailed means for governing data sharing services as done by EU's Draft Data Governance Act.

To end, we will like to draw attention to the working draft of our paper, 'Economic Governance of Data',[23] which compares European data approach with this new evolving Indian approach of 'a community's right to its data'. The paper concludes that the two approaches incorporate some important and complementing elements. Taken together these can provide the basis of a comprehensive framework of data rights and data governance. It would protect individual and collective rights of people, and those of small business actors, in the digital economy. Such a composite framework will also ensure appropriate digital industrialisation of countries. This is urgently required to escape the unsustainable situation of a global digital economy that is fast – and perhaps soon irreversibly – concentrating over two global digital poles, of US and China.

[23] https://itforchange.net/sites/default/files/1880/Economic-governance-of-data.pdf