ALG India Law Offices LLP
19, Basement, Anand Lok,
New Delhi – 110049
(T): +91.11.4152.1980
(E): ip@algindia.com

Shri. Kris Gopalakrishnan
Chairperson,
Committee of Experts
Data Governance Framework

September 13, 2020

Re: Report on Non-Personal Data Governance Framework

Dear Shri Gopalakrishnan,

ALG India Law Offices LLP ("ALG") submits these comments in response to the invitation for feedback on the Report on Non-Personal Data Governance Framework prepared by the Expert Committee (hereinafter referred to as the 'Committee') under your Chairmanship constituted vide OM No. 24(4)2019- CLES on September 13, 2019.

ALG represents several domestic and foreign companies having an interest in data governance. In the course of advising its clients, ALG has had the occasion to consider and reflect on the current legal landscape in India pertaining to data governance in general, and the aforementioned report in particular.

Our key comments and suggestions, discussed in detail in the enclosed Note, are summarized below -

1. Lack of clarity on the stakeholders with whom consultations were held by the committee and the process adopted in respect thereof.

2. The concept of collective privacy as well as provisions in respect of governance of data pertaining to collective privacy should be elaborated.

3. Data-sharing purposes encompass a broad range of activities which could be of concern from the perspective of dilution of IP rights.

4. Lack of specificity in 'social/ public/ economic benefit perspective' which is to be adopted by the Non-Personal Data Authority in deciding cases entailing refusal of data sharing requests.

5. There should be clear provisions for protection for mixed datasets.

6. If categorisation of data is done on the basis of its creation, the extent of protection of data needs to be specified.

7. Requirement of penal consequences of unauthorized usage of data.

8. Non-Personal Data need not necessarily inherit the sensitivity characteristic of the underlying Personal Data from which the Non-Personal Data is derived.

ALG India Law Offices LLP
19, Basement, Anand Lok,
New Delhi – 110049
(T): +91.11.4152.1980
(E): ip@algindia.com

9. Graduated sharing obligations should not apply to sensitive and critical data.

10. Algorithms / proprietary knowledge should not be considered for data sharing, unless voluntarily disclosed.

11. Adequate mechanisms should be put in place for protection of data collected during registration as a data business.

12. Government should consider some other sectors (other than the health sector) for the pilot project.

We appreciate the considerable effort that has gone into the Report. We recognize the time pressures and challenges and under which the committee is working, particularly in COVID-19 times. We thank you for your time and consideration of these comments.

Respectfully submitted,

Gaurav Bhalla, Partner
Sri Lekha Rayapati, Associate

For **ALG India Law Offices LLP**
19 Anand Lok – Lower Ground Floor
New Delhi – 110049
India

ALG India Law Offices LLP
19, Basement, Anand Lok,
New Delhi – 110049
(T): +91.11.4152.1980
(E): ip@algindia.com

NOTE CONTAINING ALG'S COMMENTS AND SUGGESTIONS ON THE REPORT BY THE COMMITTEE OF EXPERTS ON NON-PERSONAL DATA GOVERNANCE FRAMEWORK

1. Lack of clarity on the methodology adopted by the Committee with respect to the consultation with stakeholders.

1.1. *Observations*

1.1.1. Point 2.2.1.(i) of the Report [Methodology – Consultation with stakeholders] reads as *"As part of the deliberations the Committee met with representatives from various sectors of business (Indian and global companies) to get their views - Health, e-Commerce, Internet, Enterprise Subject matter experts, Not for Profit / think-tanks, technology service providers, etc.*

1.1.2. Point 2.2.1.(ii) of the Report reads as *"Several experts too presented their ideas / views and discussed with the Committee over meetings / video conference calls / mails."*

1.2. *ALG's Comment(s)/Recommendation(s)*

1.2.1 It is suggested that the Committee disclose the list of representatives / experts with whom consultations were held or whose ideas / views were discussed by the Committee.

2. The concept of collective privacy as well as provisions in respect of governance of data pertaining to collective privacy should be elaborated.

2.1. *Observations*

2.1.1. Point 3.8.(iv) of the Report reads as *"Collective privacy refers to possibilities of collective harm related to Non Personal Data about a group or community that may arise from inappropriate exposure or handling of such data. There remain concerns about safety of all such data in relation to the interests of the group or community about which the data is, whereby the term collective privacy is employed. For example,*

- *Data emerges about people of certain sexual orientation frequenting certain pubs / restaurants. And certain other groups of people, who are opposed to this sexual orientation, take adverse action on these pubs / restaurants. The group of people with such sexual orientation can exercise their collective privacy and ensure that such information is protected.*

- *Data emerges about people who suffer from a certain disease, which in a particular society has certain social stigma attached, and that they are centred in a particular locality in the city. In response to such data, the residents of that locality are ostracized, certain services (like delivery etc.) are denied to them. The residents of the society can take recourse to protection under collective privacy.*

- *The Committee believes that this is an emerging concept that will need to be examined and defined in detail in the future."*

2.1.2. The Report has touched upon the concept of 'collective privacy' in brief but has refrained from giving a formal definition to this concept.

2.1.3. A couple of instances of misuse of data pertaining to collective privacy have been mentioned by the Committee. The instances highlight the gravity of harm which could be caused to a community through misuse of data pertaining to collective privacy. These instances are more so particular given the Indian demographics where there is already a social stigma against various segments of the society including transgenders, people suffering from HIV, etc.

2.1.4. It is also pertinent to note how some members of the Indian society have adversely reacted towards personnel engaged in the various sectors which aid in tackling the COVID19 pandemic, particularly the members of the healthcare and law enforcement.

2.1.5. While the Report has recognised that there exists a potential concern about safety/protection of data pertaining to the interests of a group or community, it has refrained from elaborating on how to put in place a framework to ensure that there are sufficient safeguards in place to tackle this issue. Instead, the Report simply states that *'The Committee believes that this is an emerging concept that will need to be examined and defined in detail in the future'*. Leaving the formulation of a framework on protection of collective privacy for the future could result in grave misuse of data until we have sufficient safeguards put in place.

2.2. ***ALG's Comment(s)/Recommendation(s)***

2.2.1. The term 'collective privacy' should be comprehensively defined to ensure clarity on the scope of protection of data which pertains to the privacy of a community or a group. There should be clarity on how protection of 'collective privacy' differs from protection of 'community non-personal data'. If these are the same concepts, there shouldn't be the need to use multiple terminologies.

2.2.2. In addition to defining the term 'collective privacy', a framework pertaining to how such data will be collected and protected (from misuse) should be put in place (instead of leaving it for the future).

3. Data-sharing purposes encompass a broad range of activities which could be of concern from the perspective of dilution of IP rights.

3.1. ***Observations***

3.1.3. Point 7.1.(i) of the Report reads as "*Data may be requested for national security, law enforcement, legal or regulatory purposes.*".

3.1.2. Point 7.2.(i) of the Report reads as *"Data may be requested for community uses / benefits or public goods, research and innovation, for policy development, better delivery of public-services, etc."*.

3.1.3. Point 7.3.(i) of the Report reads as *"Data may be requested in order to encourage competition and provide a level playing field or encourage innovation through start-up activities*

(economic welfare purpose), or for a fair monetary consideration as part of a well-regulated data market, etc."

The aforementioned provisions of the Report provide for requests being made for the underlying data for various purposes including sovereign purpose (Point 7.1), core public interest purpose (Point 7.2) and economic purpose (Point 7.3). These purposes have been very broadly interpreted in the report to include a variety of activities which could possibly pose a risk in the form of dilution of value in such data (keeping in mind the monetary effort and labour invested by the data custodian).

3.2. ***ALG's Comment(s)/Recommendation(s)***

3.2.1 We recommend that the scope of the purposes for data sharing be reconsidered keeping in mind the situations wherein data sharing is absolutely necessary (for the larger public interest). Mandating data sharing for a broad range of activities may result in dilution/disclosure of intellectual properties of organizations, thus undermining the investment by the organization towards collection and processing of the data.

4. Lack of specificity in 'social/ public/ economic benefit perspective' to be adopted by the Non-Personal Data Authority in deciding cases entailing refusal of data sharing requests.

4.1. ***Observations***

4.1.1. Point 7.5.(iv) of the Report reads as *"If the data custodian refuses to share the request, the request is made to the Non-Personal Data Authority. The authority evaluates the request from a social/ public/ economic benefit perspective. If the request is genuine and can result in such benefits, the authority will request the data custodian to share the raw/factual data. If the authority determines that the benefits are not real, the request is denied.".*

4.1.2. The term 'social/ public/ economic benefit perspective' as envisaged under the aforementioned point is too broad and lacks specificity with respect to the considerations that the Non-Personal Data Authority would take into account while assessing a denial of data sharing request. A liberal interpretation of this term could result in a potential risk of disclosure of intellectual properties of the organizations which are mandated to share data under the 'social/ public/ economic benefit'.

4.2. ***ALG's Comment(s)/Recommendation(s)***

4.2.1. We recommend that considerations to be employed by the Non-Personal Data Authority whilst deciding cases involving refusal of data sharing requests be clearly defined keeping in mind the principle of equity.

4.2.2. Where the interest of the public at large can undoubtedly be seen/anticipated coupled with the fact that the harm caused to the data custodian due to disclosure of data is not of a very high magnitude, it is only in such scenarios that the regulatory authority elect to direct the data custodian to accede to the data sharing request.

5. There should be clear provisions for protection for mixed datasets.

ALG India Law Offices LLP
19, Basement, Anand Lok,
New Delhi – 110049
(T): +91.11.4152.1980
(E): ip@algindia.com

5.1 *Observations*

5.1.1. While elaborating on the categorisation of data in Appendix 2 (to the Report), the Report refers to the concept of a 'mixed dataset'. The Report in Point 2 of Appendix 2 explains it as – *"A mixed dataset, which represent a majority of datasets used in the data economy, consists of both personal and Non-Personal Data."*

5.1.2. While the report has stated in very simple words that a dataset consisting of both Personal and Non-Personal Data is called a 'Mixed Dataset', it does not elaborate on the framework of protection while shall be applicable for protection of non-personal forming part of mixed datasets.

5.1.3. The report gives an example of the legal framework in the European Union for protection of mixed datasets. The Report reads – *"In the European Union context, the Non-Personal Data Regulation applies to the Non-Personal Data of mixed datasets; if the Non-Personal Data part and the personal data parts are 'inextricably linked', General Data Protection Regulation apply to the whole mixed dataset."* It seems that the framework for protection of mixed datasets in the European Union is clearly demarcated such that the confusion of applicable laws on relevant sets of data is minimized. On one hand, while the Report has recognised the framework of protection of mixed datasets in the European Union, it has not elaborated on the applicable framework in India.

5.2. *ALG's Comment(s)/Recommendation(s)*

5.2.1. There does not seem to be clarity to the effect whether 'mixed datasets' shall be protected by the Non-Personal Data Protection framework which is being envisaged by the Report or through the legislation arising out of the Personal Data Protection Bill, 2019.

5.2.2. It is recommended that there should be clear cut demarcation of laws applicable to mixed datasets. This is even more necessary in case of mixed datasets wherein the non-personal and personal portions of the data are 'inextricably linked', and the process of ascertaining the applicable laws is a challenge.

6. If categorisation of data is done on the basis of its creation, the extent of protection of data needs to be specified.

6.1 *Observations*

6.1.1. The Report mentions the basis of creation of data as one of the forms of categorisation of data. In this regard, Point 3 of Appendix 2 reads – *"One approach includes four categories of data: i) provided (applications registrations, survey responses, social network postings etc.); ii) observed (cookies on a website, data from sensors etc.); iii) derived (computational scores, classification based on common attributes etc.); and iv) inferred data (scores developed using statistical, advanced analytical techniques, or AI/ML)."*

6.1.2. The Report states that such a categorisation helps in framing regulation and policy. Since the Report gives emphasis on this type of categorization, elaboration on the extent to which data could be utilized by stakeholders was also desirable.

ALG India Law Offices LLP
19, Basement, Anand Lok,
New Delhi – 110049
(T): +91.11.4152.1980
(E): ip@algindia.com

6.1.3. The Report gives the example of the framework in the European Union and reads - *"…in the European Union, the right to data portability under the GDPR would include 'provided' as well as 'observed' data. It would however exclude data 'derived' (& 'inferred') from additional processing – data that are often considered proprietary."* On similar lines, it would be extremely useful for the Government (while drafting a legislation for protection of non-personal data) if there is elaboration on the extent of right of utilization of data (if the Government elects to categorize data on the basis of its creation).

6.2 *ALG's Comment(s)/Recommendation(s)*

6.2.1. In addition to listing out the various forms in which data could be categorized, the Report should ideally have recommended a form of categorisation which seems to be most suitable, efficient and effective. In addition to this, there also needs to be elaboration on the framework of protection of data if either of the forms of categorisation of non-personal adopted is adopted.

6.2.2. Further to the above suggestion, if the Government decides to adopt the categorisation of data based on creation of data, there should be clarity on the type and the extent to which non-personal data could be used. In other words, there should be clarity whether India will adopt a similar approach to that of the European Union [(and afford the right to data portability only to data 'provided' and 'observed' (and not to 'inferred' and 'derived' data)].

7. Requirement of penal consequences of unauthorized usage of data.

7.1 ***Observations***

7.1.1. It is worth appreciating that the Committee has recommended that the data principal should also provide consent for anonymisation and usage of this anonymized data while providing consent for collection and usage of his/her personal data. It is foreseeable that there might be instances where consent was given by the data principal for collection and usage of his/her personal data, but the same was anonymized and used without the data principal knowing that it is being used for such a purpose (which the data principal might not agree to).

7.1.2. In this regard, the Report as part of Recommendation 1 recommends that – *"…the data principal should also provide consent for anonymisation and usage of this anonymized data while providing consent for collection and usage of his/her personal data."*

7.1.3. Further, the Report while discussing the checks and balances to be put in place for the data protection framework has very briefly and generally discussed the liability arising out of breach. Point 7.6(vi.) in this regard reads as - *"One reason for standards driven approach is that organisations, that comply thoroughly with the laid-down standards via annual lightweight self reported, self-audited digital compliance reports, exhibit good faith and have best-effort internal processes in-line with the best of industry standards, are to be indemnified against any vulnerability found as long as they swiftly remedy it."* As we can see, the Report does not provide for any penal actions as part of breach of liability. In fact, it provides for indemnification of the party at whose account the breach has been caused (if they are able to swiftly remedy the breach). The Report seems to have provided a long rope to the parties whose actions result in breach of protection of non-personal data.

ALG India Law Offices LLP
19, Basement, Anand Lok,
New Delhi – 110049
(T): +91.11.4152.1980
(E): ip@algindia.com

7.2. ***ALG's Comment(s)/Recommendation(s)***

7.2.1. The Non-Personal Data Protection Framework should envisage and mention specifically that there would be penal consequences in cases of unauthorized usage of non-personal data [including non-consensual anonymisation of personal data (and usage thereof)].

7.2.2. Severe penal actions in the envisaged legislation would have a much impactful deterrent effect for any person/entity attempting to anonymize personal data and/or use it without the consent of the data principal.

8. Non-Personal Data need not necessarily inherit the sensitivity characteristic of the underlying Personal Data from which the Non-Personal Data is derived.

8.1. ***Observations***

8.1.1. In most cases, it is anticipated that the non-personal data derived from the personal data would inherit the same level of sensitivity as that of the personal data as prescribed under the framework for protection of personal data viz. the Personal Data Protection Bill, 2019.

8.1.2. Point 4.5(v) of the Report mentions that *"The Committee recommends that Non-Personal Data inherits the sensitivity characteristic of the underlying Personal Data from which the Non-Personal Data is derived."*

8.1.3. While this kind of an approach makes sense for most cases, there might be some unanticipated scenarios wherein there might be requirement for the non-personal data to be afforded a higher/lesser degree of sensitivity in comparison to that of the underlying personal data.

8.2. ***ALG's Comment(s)/Recommendation(s)***

8.2.1. ALG recommends that the sensitivity characteristic of non-personal data need not be strictly as per the sensitivity characteristic of the underlying personal data. While there could be a general rule stating that the sensitivity characteristic of the non-personal data would follow that of the underlying personal data, there should be a provision for the legislature or the executive to exercise power to afford a higher/lower degree of sensitivity to the non-personal data (derived from some personal data).

9. Graduated sharing obligations should not apply to sensitive and critical data

9.1. ***Observations***

9.1.1. The Report recognises that it is necessary to provide incentives for small businesses and start-ups since symmetric data sharing obligations might not always work to promote the small businesses and start-ups to ensure a level playing field.

9.1.2. In this regard, Point 4.8(iv) reads as *"An appropriate Non-Personal Data framework legislation, while providing community data rights will also lay down principles and guidelines for various incentives for data custodians, respective data privileges, compensations, where needed, the nature of the required, well-regulated data markets, and so on."* It further goes on to

state that *"The framework law will also provide means to specifically protect and promote the interests of small Indian companies and startups. Symmetric data sharing obligations equally on all data businesses may not always work for small businesses, and may even be to their detriment. Provisions like threshold size for data sharing, and graduated sharing obligations, may be considered."*

9.1.3. While the concept of graduated data sharing seems logical for the betterment of small scale Indian entities, there needs to be sufficient safeguards in place such that all data is not shared in an unhindered manner with small entities and start-ups, which might lead to the data being adversely utilized by such entities.

9.2. ***ALG's Comment(s)/Recommendation(s)***

9.2.1. The Non-Personal Data Protection framework should specify that while graduated data sharing obligations could be there to provide an incentive to small Indian companies and startups, the graduation should not be such that 'sensitive' and 'critical' data is shared without any restriction with small Indian companies and startups. There should be a specific provision ensuring that the graduated data sharing obligations only extend to general non-personal data.

9.2.2. Further, there should be a clear definition identifying the threshold (in monetary terms or in any other appropriate form) of companies that would qualify as 'small companies'. The envisaged framework could also look at adopting the current category of MSME's to be qualified as small companies (if such a categorisation is found suitable).

10. Algorithms / proprietary knowledge should not be considered for data sharing, unless voluntarily disclosed

10.1 ***Observations***

10.1.1. Copyrighted data or any data which would fall under the purview of trade secret should not be covered for protection under non-personal data.

10.1.1. Point 5.4(iii) of the Report states - *"Algorithms / proprietary knowledge may not be considered for data sharing."* The usage of the word 'may' in this Point results in ambiguity on the subject on whether or not Algorithms / proprietary knowledge (some of which could be the subject matter of copyright and trade secrets) could be disclosed as non-personal data.

10.2. ***ALG's Comment(s)/Recommendation(s)***

10.2.1. The disclosure of non-personal data of the data principal in the form of Algorithms / proprietary knowledge should not be allowed, unless express consent to this effect is obtained from the data principal.

10.2.2. Public disclosure of such data without the consent of the data principal will lead to the effort (in the form of time and money) of the data principal being non-utilized and would not serve as a motivating factor for the Indian entities to invest in innovation and research. This could serve as a deterrent for individuals and companies to invest in R&D when they would be under constant threat that their algorithms / proprietary knowledge could be made public.

ALG India Law Offices LLP
19, Basement, Anand Lok,
New Delhi – 110049
(T): +91.11.4152.1980
(E): ip@algindia.com

11. Adequate mechanisms should be put in place for protection of data collected during registration as a data business.

11.1. *Observations*

11.1.1. The Report provides for collection of various information while registering a business as a data business.

11.1.2. Point 6.2(iii) of the Report states that – *"Initial registration would require a business ID (or country code and country business ID), digital platform/business name(s), associated brand names, rough data traffic and cumulative data collected in terms of number of users, records and data. Also needs to be stated is, the nature of data business, and kinds of data collection, aggregation, processing, uses, selling, data-based services developed etc."*

11.1.3. Further, Point 6.3(ii) of the Report reads as – *"Every Data Business must declare what they do and what data they collect, process and use, in which manner, and for what purposes (like disclosure of data elements collected, where data is stored, standards adopted to store and secure data, nature of data processing and data services provided)…"*

11.1.4. The information sought for registering a business as a data business, particularly 'kinds of processing, uses, sale and data-based services developed, etc.' could lead to the entity disclosing its manner of functioning, which could be prejudicial to its interests if such information reaches its competitors. Further, since the list is non-exhaustive, the entities cannot actually predict the information that they might be asked to disclose for a registration as a data business.

11.2. *ALG's Comment(s)/Recommendation(s)*

11.2.1. The framework should provide for sufficient safeguards to ensure that the information shared by entities for registration as data businesses is safely secured, and does not reach in the hands of the competitors of the registrant of the data business. The framework for protection of non-personal data should also mention an exhaustive list of information that will be required for registration of an entity as a data business. This will enable entities to manage the threshold of data collection to decide whether they want to fall in the category of a business (and be subject to such disclosure of information).

12. Government should consider a sector (other than the health sector) for the pilot project.

12.1 *Observations*

12.1.1. In Point 7.2(iv), the Report recommends that the health sector could be considered as a pilot-use case for Non-Personal Data Governance Framework. To elaborate on this, the Report states – *"large anonymised data sets of health data, could lend community level insights into diseases, epidemics, and community genetics – leading to better tailored health solutions for the community. Accordingly, the Committee considered health as a pilot use-case for the Non-Personal Data governance framework".*

12.1.2. While there is no doubt about the anticipated benefits arising out of processing of information contained in non-personal data pertaining to the health sector, its privacy concerns

ALG India Law Offices LLP
19, Basement, Anand Lok,
New Delhi – 110049
(T): +91.11.4152.1980
(E): ip@algindia.com

and the detrimental effect it would have on individual(s) and communities as a whole cannot also be undermined.

12.1.3. Also, since there might be concerns whether sufficient safety standards are in place to ensure anonymisation of data or not, the healthcare sector might not be a suitable sector for conducting the pilot project.

12.2. ***ALG's comments/recommendations***

12.2.1. In the Indian socio-economic scenario, there are frequent instances of discrimination with person(s) (and communities as a whole) suffering from diseases/ailments by the public at large. The most recent and prominent instance of this could be seen in the ongoing pandemic where there were numerous instances of discrimination against not only people suffering from COVID19 but also healthcare and other social welfare workers involved in the people suffering from COVID19.

12.2.2. The concerns of social detrimental effects coupled with anonymisation concerns of such data, the health sector would not be the appropriate sector for the pilot project for non-personal data governance framework.

12.2.3. A sector (such as the FMCG sector) should be chosen for the pilot project since it may contain less sensitive personal data and the chances of harm being caused by misuse of such data are much lesser.