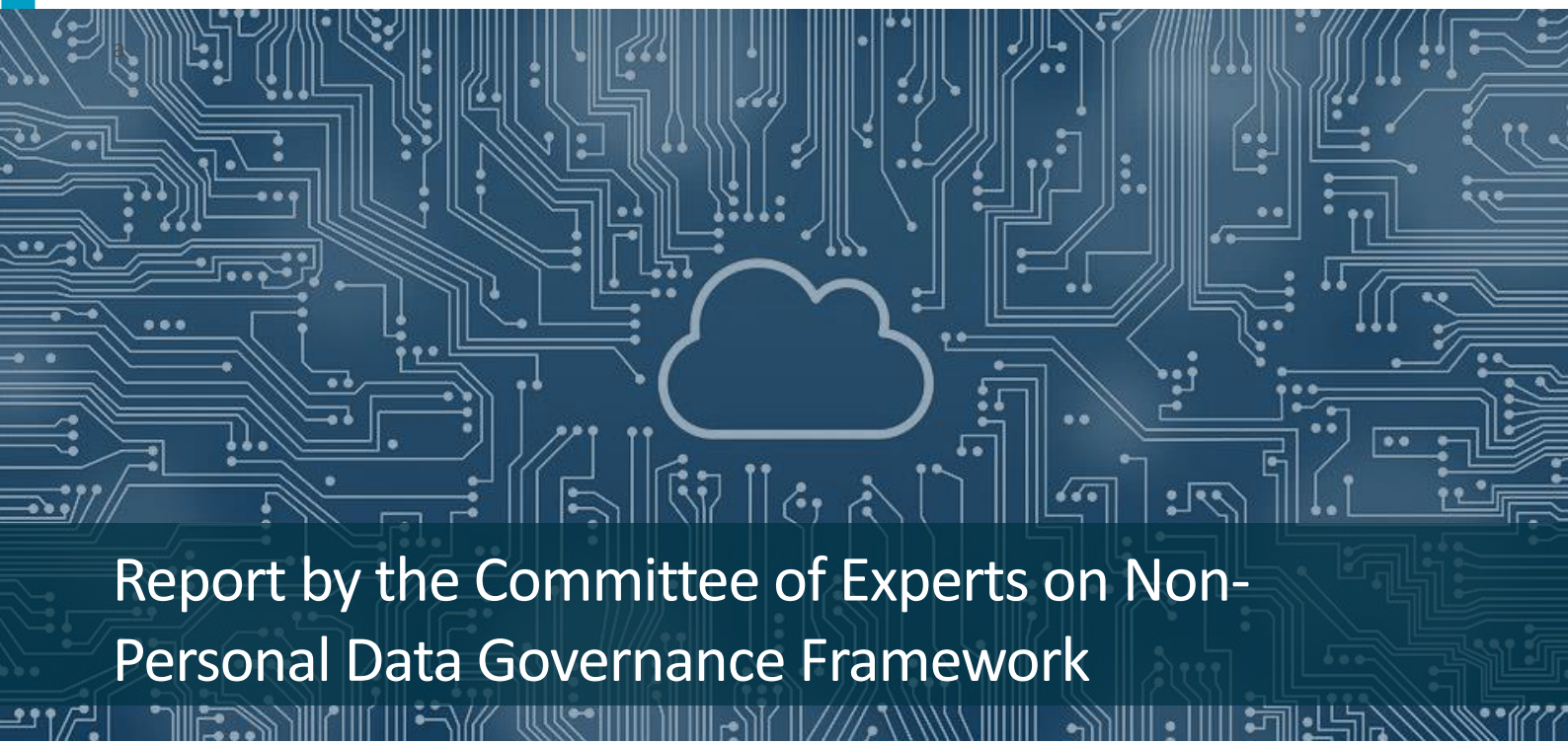
ANALYSIS

TRILEGAL



Report by the Committee of Experts on Non-Personal Data Governance Framework

17 July 2020

This analysis covers:

The Ministry of Electronics and Information Technology (**MEITY**) constituted a committee of experts (**Committee**) in September 2019, to devise a framework for regulation of non-personal data (**NPD**). The Committee released its report (**Report**) on 12 July 2020 for public consultation/feedback, and has set a deadline of 13 August 2020 for comments from the public. The Committee has proposed the introduction of legislation governing NPD (**NPD Legislation**), to be enforced by an NPD authority (**NPDA**) and lays down key principles to be incorporated in the NPD Legislation. This is a distinctive move given that at present most other nations, including the European Union, have focused primarily on regulating the flow of NPD between countries. The NPD Legislation would be one of the first to govern the collection and processing of NPD within a jurisdiction.

We have briefly analysed below the key principles and recommendations of the Committee.

The Report recommends the introduction of a new legislation and regulatory authority to govern non-personal data. The proposed legislation, if modelled along the lines of the Report, is likely to significantly impact the manner in which businesses use, process, store and share non-personal data.

1 Rationale for an NPD Framework

The Report proposes the regulation of NPD on the primary ground that the economic and social value of data is concentrated in the hands of a few companies who have become monopolies in the absence of any regulation. The introduction of an NPD framework is aimed at catalysing data companies such that the welfare of all relevant stakeholders is maximised. In recognising this value of NPD, the Report states that the lack of regulation denies both individual citizens and communities the benefits that they can derive from such data, and also gives rise to a risk of collective harm from misuse. The Committee believes that a balanced regulation for NPD will spur improved service delivery, innovation and research by both public and private sector entities, while specifically benefitting the India start-up ecosystem.

2 Definition and classification of NPD

The Committee defines NPD as data that either (i) never related to an identified or identifiable natural person; or (ii) is sourced from personal data (**PD**)[1], as defined under the Personal Data Protection Bill, 2019 (**PDP Bill**), i.e. it is data which was initially personal but was later aggregated and made anonymous. However, the Committee has recognised that anonymised data has the potential to be re-identified. Certain data aggregation processes may also result in data being continuously anonymised and re-identified, and it is therefore unclear how such data will be treated under the proposed framework.

Based on this definition, the Committee has broadly categorised NPD into the following:

a. *Public NPD* - collected or generated by the government or in the course of publicly funded work and excluding data that is treated confidentially under any law, such as anonymised land records and vehicle registration data.

[1] Personal Data is defined under Section 3(28) of the PDP Bill as data about or relating to a natural person who is directly or indirectly identifiable, having regard to any characteristic, trait, attribute or any other feature of the identity of such natural person, whether online or offline, or any combination of such features with any other information, including any inference drawn from such data for the purpose of profiling.

b. *Private NPD* - collected or produced by persons or entities other than the governments, the source or subject of which relates to assets and processes that are privately-owned, and includes derived/inferred data that result from private effort, insights involving application of algorithms or proprietary knowledge, and data included in a global dataset pertaining to non-Indian which is collected in foreign jurisdictions. This would therefore include any and all data that is collected by private entities that it is not Community NPD.

c. *Community NPD* - whose source or subject is a community, i.e. a group of people that are bound by common interests and purposes and involved in social/economic interactions, including entirely virtual communities. This may include data collected by the municipal corporations, public electric utilities, and private players like telecom, e-commerce, and ride-hailing companies.

Community NPD is a concept introduced by the Committee that does not presently find place in equivalent global legislations.

Community NPD also specifically excludes Private NPD. The Report clarifies that raw data collected by private entities from a community would constitute Community NPD, while inferred or derived data would constitute Private NPD. However, Community NPD appears to overlap both with Public NPD and Private NPD. For instance, the Report states that Community NPD includes data collected by municipal bodies, which would also clearly fall within the definition of Public NPD. Similarly, many NPD data sets relating to group of customers including from virtual communities (such as social media users) form part of Community NPD. The baseline definition of these three fundamental concepts itself is confusing under the report.

Further, the definition of Private NPD includes data in a global dataset pertaining to non-Indians which is collected in foreign jurisdictions. Inclusion of non-Indian data does not align with the stated objectives of the Report.

3 Sensitive NPD

Given that sensitive or personal information can lead to collective privacy harms, even in the form of NPD, the Committee has defined a new concept of Sensitive NPD, which is NPD that may relate to (i) national security or strategic interests such as vital infrastructure; (ii) business sensitive or confidential information; or (iii) anonymised data, that bears a risk of re-identification.

With respect to anonymised data, the Committee has recommended that all NPD that is derived from sensitive personal data (SPD)[2], as defined under the PDP Bill, should inherit the sensitivity of the underlying SPD. However, unlike the increased compliances for SPD under the PDP Bill, the Committee has provided no information as to the additional measures to be taken to protect Sensitive NPD, aside from certain cross-border transfer restrictions.

Further, such a principle is only beneficial in cases where the underlying data can be clearly categorised as SPD, such as health data. In this respect, the classification of NPD as sensitive should not solely rely on the presence of underlying SPD, given that even NPD that has no underlying SPD datasets may be also be considered sensitive. For instance, a dataset containing information regarding potential COVID-19 risk areas in a city could be created by layering non-sensitive datasets of the address and travel history of persons in a

[2] SPD is defined under Section 2(36) of the PDP Bill to include personal data that may relate to, reveal or constitute (i) financial data; (ii) health data; (iii) official identifier; (iv) sex life; (v) sexual orientation; (vi) biometric data; (vii) genetic data; (viii) transgender status; (ix) intersex status; (x) caste or tribe; (xi) religious or political belief or affiliation.

city. Such a dataset is likely to be considered Sensitive NPD even though neither of the underlying datasets constitute SPD.

4 Consent for Anonymised Data

Recognising the risk of anonymised data being de-anonymised, the Committee has sought to protect individuals from any harm arising out of such re-identification. The Committee has recommended that appropriate standards of anonymisation be defined to prevent or minimize the risks of re-identification, and has laid down recommended anonymisation processes that may be followed. It is important that such standards be harmonised with those notified under the PDP Bill, to ensure conformity between the two regimes. Further, the Committee has recommended that any PD that is being anonymised should continue to be viewed as NPD of the provider of such information. The individual should therefore be required to provide consent for the anonymisation of the PD and the subsequent use of the resulting anonymised data. This consent can be obtained at the time of collection of the PD under the PDP Bill.

5 Key Stakeholders

The Report lists the following roles for potential players within the NPD framework:

a. *Data Principal* - In case of Public NPD and Private NPD, this is the person (individuals, companies, communities) to whom the data relates. In the case of Community NPD, the community that is the source of the NPD would be the Data Principal. This is similar to categorisation of a data principal under the PDP Bill, in relation to PD, with Data Principals being allowed to exercise significant control and economic rights over their NPD.

b. Data Custodian - This is the person who undertakes collection, storage, processing and use of NPD. Data Custodians may be public or private sector entities who process NPD such as government ministries, telecom companies or e-commerce entities. Data Custodians must comply with requirements under the NPD Legislation, such as adopting prescribed anonymisation standards. NPD must be used by Data Custodians in a manner that is in the 'best interest' of the Data Principal. They have a 'duty of care' to the individual or community from which NPD has been collected. This principle is similar to that of a data fiduciary under the PDP Bill, which lays down specific obligations to be undertaken by the data fiduciary with respect to the data rights of the Data Principal. However, the standard for ensuring the 'best interest' of the Data Principal is presently unclear, and the Committee has recommended that they be detailed in the NPD Legislation. They have indicated that such standards may relate to anonymisation techniques and graduated data sharing.

c. Data Trustee - This is the person through which a community exercises its data rights and who takes action to protect the community against any collective harm arising from the use of Community NPD. In most instances, the Data Trustee will be the closest and most appropriate representative body for a community, and may be a government agency at any level (such as the Ministry of Health for data on diabetes in India). However, it could also be citizens' groups (such as residents' welfare associations for local data), or civil society organisations. However, there is no clarity provided as to how a Data Trustee would be identified, the eligibility criteria for such an entity, or whether the community data principals play a role in identifying the Data Trustee, and this is to be provided under the NPD Legislation.

   This may result in the existence of multiple Data Trustees having rights over the same NPD, and to account for this the Committee has proposed the introduction of a digital NPD Policy Switch (Policy Switch) (as

discussed in Section 12), which is a tool to manage NPD regulations. This mechanism seems to suggest that Data Trustees can issue regulatory guidelines, whereas in all other sections of the Report they are merely mentioned as having the right to make recommendations to the NPDA with respect to Data Custodians. This could lead to multiplicity of regulations given that there could be different Data Trustees for the same NPD.

In order to protect communities against collective harm, Data Trustees have the power to make recommendations to the NPDA to enforce obligations of Data Custodians with respect to data practices under the NPD Legislation, such as sharing of data reporting mechanisms, transparency, and enforcement of safeguards. The obligations of the NPDA to consider or enforce such recommendations under the NPD Legislation would be critical in determining the extent of the rights that Data Trustees would have over Data Custodians.

d. *Data Trust* - This is an institutional structure bound by rules for handling a specific set of NPD. Such trusts may hold NPD which may be voluntarily shared by Data Custodians, or mandatorily shared NPD on the basis of orders from the government or Data Trustees (as described below in Section 8). However, the Committee has provided very little insight as to how Data Trusts will function, including how such trusts will be constituted, who determines its members, and its role in the NPD ecosystem.

6 Ownership of NPD

Based on the categories of NPDs described above, the Committee has made recommendations regarding the ownership of NPD by the relevant stakeholders. On the grounds that Public NPD is derived from public efforts, such data should be akin to a national resource, and ownership would therefore lie with the state.

The Committee has recognised that there may be an overlap in ownership of data with respect to Community NPD and Private NPD and has therefore introduced the principles of 'best interest' and 'beneficial interest'. In case of NPD that is derived from PD of an individual, the provider of the PD will continue to be the Data Principal for the NPD, and the NPD should be utilised in the 'best interest' of that individual. It is unclear what the parameters of 'best interest' are in the case of Data Custodians, and whether this will restrict the ways in which NPD can be used. This could have significant implications since most companies often use NPD for their benefits that may not specifically benefit Data Principal.

Similarly, for Community NPD, while the beneficial ownership belongs with the community, the rights over such data are to be exercised by the Data Trustee in their 'beneficial interest'. The principle of 'beneficial interest' necessarily requires decisions to be taken on the basis of the needs of the community as a whole, and therefore may be used for purposes that certain members of the community may not agree with. The Committee has not fleshed out the decision-making processes of the Data Trustee, and how individual rights of NPD will be balanced against the rights of the community. Further, there is no clarity on how Data Trustees will determine 'beneficial interest' of the community collectively and whether this will be done in consultation with members of a community, with no safeguards being provided in this regard.

7 Applicability

The Committee has referred to the fact that NPD collected about people in India and collected in India would be subject to Indian laws and regulations. This would mean that all Indian Data Custodians, including those collecting NPD of foreign individuals, and all foreign Data Custodians collecting NPD of Indians would be

subject to these regulations. Therefore, the presence of Indian NPD in a dataset will invoke applicability of the NPD Legislation, even for datasets created by foreign entities which largely contain NPD of foreign individuals.

Similarly, in light of the ambiguity around the scope of the term 'community', and given that Community NPD includes virtual communities, it is unclear whether foreign members of such communities will be given rights under the NPD Legislation, and how they will exercise such rights. There also appears to be no explicit restriction on foreign Data Trustees, as long as they can show that they are the closest and most appropriate representative body for a community which may have Indian members. For instance, this may be the case for purely virtual communities, such as international social media and e-commerce platforms, which may have Indian users. It is unclear whether the 'beneficial interest' obligation would be restricted only to Indian users of the community in such a case. This also becomes significant given that foreign Data Trustees may then be given a right to make recommendations to the NPDA.

8 Data Businesses

The Committee has recommended that private and public sector entities who collect NPD be required to register as a Data Business, subject to meeting certain criteria as may be determined by the NPDA. For entities who do not meet this threshold, this requirement will be voluntary. The Committee has recommended that this registration be a one-time lightweight and fully digital process, wherein entities have to share details regarding what they do and what data they collect, process and use, in which manner, and for what purposes. To streamline the process, it is recommended that these disclosures be made along with those relating to PD under the PDP Bill, if applicable.

In a first-of-its-kind move, the Committee has suggested that if data traffic of a Data Business exceeds certain limits, then the metadata regarding the NPD being collected and processed by Data Businesses should be made publicly accessible. This will allow citizens, organisations and government agencies to identify opportunities to use such NPD, either independently or in combination with each other, to innovate. Data Businesses can be requested to share the underlying NPD with such persons in accordance with a specific data sharing procedure (as discussed in Section 10). Given the lack of precedent for the success of such a requirement, it is important to ensure that such a mechanism does not disincentivise Data Businesses from processing NPD, as this would defeat the Committee's aim of leveraging NPD collected by data-intensive businesses for innovation, and social and economic development.

9 Data Sharing Purposes

The Committee has identified three purposes for which Private, Community and Public NPD may be shared with citizens, companies, policy and research organisations and the government:

a. sovereign purposes including data requests for national security, law enforcement, legal or regulatory purposes;

b. public interest purposes including research, innovation, policy development and delivery of public services, and

c. economic purposes such as encouraging competition and innovation.

While businesses may use the shared NPD for private purposes, Data Trustees and the government may request NPD from businesses and make it available to relevant parties.

Notably, algorithms and other proprietary knowledge have been specifically excluded from the obligation to share. In this regard, the Draft E-commerce Policy published by the Department for Promotion of Industry and Internal Trade (Draft Policy) specifies that the government may request disclosure of source codes and algorithms from e-commerce entities to protect against digitally induced biases. Consultations with other sectoral regulators would be required to avoid such discrepancies.

The Report encourages Indian businesses to use data analytics, machine learning and artificial intelligence to leverage the shared metadata and underlying NPD to create new business models by identifying unique digital solutions. The Committee has also recognised that a data sharing framework is likely to give rise to data marketplaces where trusted intermediaries enable transactions between a data supplier and a data user for a remuneration. The existence of an open data sharing framework will increase transparency in the nature and extent of data processing undertaken by companies.

Permitting government access to NPD is a common legislative provision across jurisdictions. Particularly in the EU, the government authority may access NPD processed by an entity incorporated within its jurisdiction even if the processing takes place overseas. It is unclear from the Report whether the Committee recommends a similar reach for the Indian government. The extent of government access to NPD must be clarified in the NPD Legislation. The Draft Policy identifies data storage service as a potential key export for India. While this may incentivise foreign companies to store their data in Indian data centres, data sharing obligations imposed by the NPD Legislation may disincentivise such a move. The NPD Legislation should clarify whether Indian stakeholders will have access to NPD of foreign companies that may be stored in India. Notably, there is no clear exemption for data processors, which could imply that business process outsourcing and cloud service providers may also be covered under the NPD framework.

10 Data Sharing Mechanisms

The Committee has recommended the establishment of appropriate NPD sharing mechanisms. Particularly with respect to sharing private NPD, the Committee recommends mandatory sharing of sufficiently anonymised raw/factual NPD at no cost. When there is a value addition, the private companies can charge for the NPD that it shares. The charges would depend on the level of the value addition. For non-trivial value additions, NPD sharing may be mandated but subject to fair, reasonable and non-discriminatory (FRAND) based remuneration, for an increased value addition the charges will be determined by market forces, and for 'high value add' NPD, the private companies will have full discretion.

The Committee's recommendation that private companies are not entitled to a fee for raw NPD may disincentivise companies from collecting and storing NPD, which would defeat the purpose of such a sharing mechanism. While the Committee recommends affixing costs proportionate to the value addition, there is a lack of clarity on what would constitute 'non-trivial', 'increased' and 'high' value additions. Presumably, the restrictions on cost of NPD are aimed at easing access to NPD for Indian companies. Given this context, the NPD Legislation needs to clarify whether Indian companies may sell NPD to foreign companies without restrictions. Most significantly, by regulating the charges for NPD, this recommendation restricts the ability of private businesses to enter into commercial contracts for the sharing of data on their own terms.

The Committee recommends that NPD sharing may be initiated by a data sharing request. If a request is denied, the aggrieved business may approach the NPDA for a final decision. If the NPDA determines that the request is genuine and can result in social, public or economic benefits, it can mandate the business to share the raw or factual NPD. Therefore, the Report empowers the NPDA to override a business' discretion and

mandate the sharing of NPD in certain circumstances. Further, NPD Legislation must provide detailed guidelines for framing data sharing requests to ensure that these are reasoned and specific, and do not impose onerous obligations on private companies to routinely transfer large quantities of NPD.

11 Checks and Balances

The Committee recommends the following checks and balances with respect to NPD sharing:

a. Localisation - The Report replicates the localisation requirements under the PDP Bill and states that general NPD may be stored and processed globally however Sensitive NPD can be transferred overseas but must be stored in India, and Critical NPD can only be stored and processed in India. The NPD Legislation will continue to apply to the regulation of community NPD and public NPD transferred overseas. Businesses that transfer community NPD or public NPD overseas will be responsible for complying with data sharing requirements.

b. Contract - Cloud provider and Data Business should contract to comply with the NPD Legislation.

c. Probing - Automated tools and registered experts will probe the cloud and the shared NPD for vulnerabilities and compliances.

d. Liability - Organisations that comply with the prescribed standards will be indemnified against any vulnerability so long as they remedy it swiftly. Standards applicable to companies processing NPD and specific compliances for Data Business will be prescribed by the NPDA.

12 Non-Personal Data Authority

To regulate the collection, processing, storage and sharing of NPD, the Committee recommends the creation of a separate NPDA. The constitution of the NPDA remains to be fleshed out but the Committee has recommended that the NPDA should have some members with relevant industry experience. While the Data Protection Authority (DPA) under the PDP Bill is meant to protect PD, the NPDA is meant to unlock value in the NPD. The NPDA is expected to work harmoniously with the DPA as well as other sectoral regulators such the Competition Commission of India.

The Committee suggests that the NPDA should play an enabling role as well as an enforcing role. In its enabling role, the NPDA should ensure that NPD is available for various social, public and economic purposes. Particularly, the NPDA should enable legitimate NPD sharing requests, regulate and supervise NPD sharing agreements between relevant stakeholders and supervise the market for NPD. In its enforcing role, the NPDA should administer the provisions of the proposed NPD Legislation. This will include regulating Data Businesses, mandating the sharing of NPD in certain circumstances, setting standards and certifying frameworks, including for NPD sharing, NPD safety and anonymisation of PD.

The Committee has also proposed the introduction of a Policy Switch, which will be a single digital clearing house for regulatory management of NPD, and the performance of encoding, rationalisation (to prevent contradiction), implementation and enforcement functions. While the Report suggests that the Policy Switch will be subject to regulatory guidelines issued by Data Trustees, it recommends wider functions for it. The NPD Legislation should set out the precise scope, applicability and functions of the Policy Switch. Further, the NPD Legislation would need to clarify the role of the NPDA with respect to the Policy Switch since the Report suggests that it will be housed within the NPDA.

Although the Committee has distinguished the intended roles of the DPA and NPDA, there is likely to be a significant overlap in the activities carried out by them given the close inter-relation between PD and NPD. For instance, both regulators will exercise their oversight on Data Businesses, which are required to make NPD disclosures along with disclosures on PD under the PDP Bill. Further, entities classified as Data Businesses may overlap with entities classified as Significant Data Fiduciaries under the PDP Bill. There is also a likelihood of anonymised data being de-anonymised or re-identified resulting in privacy concerns around the identification of an individual, which will invite action from both regulators. Therefore, it will be difficult to fully divorce the regulation of PD and NPD. In addition to the DPA and NPDA, the Draft Policy also suggests an e-commerce regulator with powers to regulate e-commerce data. The existence of multiple regulators will require a great deal of co-ordination and is likely to result in confusion and delays. Establishing a sole regulator will eliminate this and reduce the compliance requirements for businesses.

13 Technology Architecture

In addition to the recommendations, the Committee has also suggested some guiding principles for a technology architecture that will enable the implementation of the NPD Legislation. These include the Policy Switch as well as the following:

a. *Mechanisms for accessing data* - All shareable NPD should have a Representation State Transfer Application Programme Interface (REST API), which is a software architecture that will enable stakeholders to request and obtain data digitally and maintain a log of all requests for NPD.

b. *Distribution for data security* - NPD should be stored in a distributed format within Secure NPD Clouds to avoid single point leakage, data corruption and data loss. Despite the distributed storage, it is important to manage the NPD in a coordinated manner.

c. *Data exchanges* - The Committee has suggested a standardised data exchange approach wherein all collated NPD should be available on a data exchange and accessible to stakeholders in a standardised and usable format. NPD Legislation should clarify the scope of 'collated NPD' and provide guidelines on the creation, operation and management of data exchanges.

d. *Effective anonymisation* - The Committee has suggested nine industry tools including differential privacy algorithm and homomorphic encryption as mechanisms to prevent de-anonymisation. These algorithms introduce restrictions that minimise the chance of re-identification of anonymised data through linkages or other means. The Committee has also recommended that academia and industry must jointly evolve best-of-breed algorithms for effective anonymisation. The NPDA will identify the minimum threshold of anonymity based on acceptable standards.

The guidelines proposed by the Committee for the storage, maintenance and sharing of NPD would require the incorporation of standards at a system level. In order to comply with it, companies will need to restructure their existing data architecture.