January 2021

Access Now's position on the Data Governance Act

Introduction

Access Now welcomes the opportunity to provide comments on the European Commission's proposal for a Data Governance Act (DGA) to inform the upcoming legislative debate in the European Parliament and in Council.

With the introduction of the DGA, we note a shift in the European Commission's approach to the governance of personal data. With the General Data Protection Regulation (GDPR) and the ePrivacy Directive, the European Union's approach was rooted in core privacy and data protection principles, including data minimisation and the protection of people's fundamental rights. The DGA provides a shift towards a data-driven economy and opens the door to the monetisation of certain practices. This shift and the data sharing model presented by the Commission was criticised in an unpublished report of the European Commission's Regulatory Scrutiny Board (RSB) when it issued a "negative" opinion on an impact assessment accompanying the proposed Data Governance Act. Even if the board later changed its opinion to "positive with reservations" following changes to the draft law, some of its criticisms "go to the heart of the new law, including saying the impact assessment "does not explain the problem clearly enough and why the EU should promote a new model for data sharing"" according to media reporting.[1]

With our comments, we seek to ensure that Europe's legacy on data protection is upheld within the DGA for the development of a sustainable and user-centric data economy. We do note that the proposal from the European Commission seeks to introduce a number of safeguards for the rights to privacy and data protection, although these need to be complemented and clarified.

For legal certainty and to maintain a high level of personal data protection in the EU, our preferred approach would be for the European legislator to exclude personal data from the scope of the DGA. This limitation would avoid overlap and possible contradictions with the GDPR. Nonetheless, we provide below comments and recommendations with the view of protecting the European data protection acquis as much as possible, strengthening safeguards for users, and increasing legal certainty for entities seeking to use the measures provided for under the DGA.

[1] Clark,S. Revealed: the Data Governance Act's teething problems, Global Data Review (2021). https://globaldatareview.com/data-localisation/revealed-the-data-governance-acts-teething-problems

Our comments focus on the following issues:

1. Clarifying the scope of the DGA and definitions
2. Addressing the risks of overlaps and contradictions with the GDPR in the area of "data altruism" and "data re-use"
3. Preventing the monetisation of data processes
4. Establishing Data Protection Authorities as main supervisory authorities under the DGA

1. Clarifying the scope of the DGA and definitions

Personal data and non-personal data

The DGA covers both personal data and non-personal data, both defined based on the definition of personal data established under the GDPR. In fact, the DGA defines non-personal data by exclusion, meaning that it covers anything that is not personal data under the GDPR.

In practice, however, non-personal data can be many different things that should be distinguished and not be taken as one category. Non-personal data can for instance include "anonymised personal data" and "non-human data".[2]

- **Anonymised personal data** is data that originally belonged to, related to, or identified a data subject but has subsequently been anonymised.
- **Non-human data** refers to information that does not come from any data subjects. This could include climate or weather data, supply chain data or data from industrial machines not linked to an individual

It is common that datasets include many different types of data - including personal data, anonymised personal data, and non-human data - or a combination of these. This overlap means that it is nearly impossible to apply safeguards for one and not the others. While increased safeguards and requirements apply to the use of personal data, the same level of protection should be given to personal data which has been anonymised or pseudonymised as it will always be vulnerable to re-identification, and should thus not be treated in the same manner as "non-human non-personal data." Moreover, consideration must be given to the fact that some non-personal data may link back to personal data or allow for personal data to be inferred. This is particularly important as the DGA particularly seeks to develop data space in the area of health and finances, among others, which indicate that sensitive data may be processed.

Definitions and references within the DGA should be updated to reflect this reality.

---

[2] Singh, A. ; Raghavan, M; Chugh, B; & Prasad, S. The Contours of Public Policy for Non-Personal Data Flows in India, *Dvara Research* (2019). *https://www.dvara.com/blog/2019/09/24/the-contours-of-public-policy-for-non-personal-data-flows-in-india/* and Marda, V. Non-personal data: the case of the Indian Data Protection Bill, definitions and assumptions, Ada Lovelace Institute (2020). https://www.adalovelaceinstitute.org/blog/non-personal-data-indian-data-protection-bill/

Data altruism and altruistic purposes

The DGA defines data altruism as follows in Article 2 (10):

"data altruism' means the consent by data subjects to process personal data pertaining to them, or permissions of other data holders to allow the use of their non-personal data without seeking a reward, for purposes of general interest, such as scientific research purposes or improving public services".

The definition allows "data holders" to authorise the use of personal data, without asking the permission of data subjects. This should be avoided as it limits people's ability to control their information and is therefore contrary to the objective of the GDPR. The definition further indicates that data will be used for "purposes of general interest" without clearly defining what these may be. While both the definition and recital 35 provide examples of what these purposes may be, it does not provide clarity as to why our information may be used. A closed list of clear purposes should be provided in the DGA to provide legal certainty and ensure compliance with the principle of foreseeability.

Beyond clarification in this definition, we note that data subjects may already consent to use of their personal data for research and scientific purposes under the GDPR. While Data Protection Authorities (DPAs) may need to provide additional guidance on how to operationalise this possibility, it is perhaps not necessary for the DGA to provide a similar avenue.

Which entities can re-use data?

While Article 1.1 of the DGA indicates that only "certain categories of data held by public sector bodies" can be re-used, it does not limit which entities may request this data in order to re-use it and why. This should be clarified by setting a closed list of clear purposes for which data re-use may be authorised.

2. Addressing the risks of overlaps and contradictions with the GDPR in the area of "data altruism" and "data re-use"

Data altruism

The DGA provides measures on "data altruism", largely governed by the definition and requirements for consent defined under the GDPR, although it goes beyond these rules. We noted above in our comment on the definition of data altruism that the DGA would allow for entities holding personal data to authorise its re-use without seeking the consent of the data subject.

The same issue appears in Recital 35 which further indicates that:

"There is a strong potential in the use of data made available voluntarily by data subjects based on their consent or, where it concerns non-personal data, made available by legal persons, for purposes of general interest."

This recital should be clarified that consent is only valid if it was given to relate to a specific purpose. Additionally, the reference to "non-personal data" should be updated to refer to "non-human personal data".

Recital 36 introduces additional concepts that directly contradict the purpose limitation principles and other requirements under the GDPR that require the data controller and processor to identify clear purposes for the processing of data:

"Data subjects in this respect would consent to specific purposes of data processing, but could also ***consent to data processing in certain areas of research or parts of research projects as it is often not possible to fully identify the purpose of personal data*** *processing for scientific research purposes at the time of data collection." (emphasis added)*

This section should be deleted.

Articles 18 and 19 provide for transparency obligations and applicable safeguards related to data altruism practices. It should be clarified that these requirements come on top of the GDPR obligations for controllers and processors to provide information to data subjects as well as all data protection rights. In addition, it should be clarified that entities seeking to process data under the rules provided by the DGA should appoint a data protection officer.

Finally, Recitals 39, 41 and 42 and Article 22 proposes that the European Commission, helped by the European Innovation Board, will be tasked with drawing a "data altruism consent form" to be filled in by data subjects to allow for the use of their information. Given the application of the consent rules under the GDPR, the need to ensure the application of data subjects' rights under Chapter II of the GDPR, and that this activity is related to data processing, it is rather Data Protection Authorities that should be tasked with developing such a form, with the help of the European Commission. The DGA cannot negate the responsibility and duties of DPAs to ensure the application of the GDPR.

Re-use of data

The DGA puts forward a series of measures for the re-use of data held by public sector bodies. In itself, this re-use is similar to what is defined as "further processing" under the GDPR. It is crucial to ensure that the DGA does not create overlapping rules with the GDPR but that it also does not negate the application of the safeguards provided under Article 6.4 of the GDPR in this scenario. In fact, to allow for the re-use of data, unless the data subject consents to a new specific purpose, the data controller should "in order to ascertain whether processing for another purpose is compatible with the purpose for which the personal data are initially collected, take into account, inter alia:"

(a) *any link between the purposes for which the personal data have been collected and the purposes of the intended further processing;*
(b) the context in which the personal data have been collected, in particular regarding the relationship between data subjects and the controller;
(c) the nature of the personal data, in particular whether special categories of personal data are processed, pursuant to Article 9, or whether personal data related to criminal convictions and offences are processed, pursuant to Article 10;
(d) the possible consequences of the intended further processing for data subjects;
(e) the existence of appropriate safeguards, which may include encryption or pseudonymisation.

The DGA should make a reference to these conditions.

Article 5 of the DGA sets the conditions that should be fulfilled for the re-use to be authorised by public sector bodies. However, similarly as in the area of data altruism, it appears that public bodies could decide on this re-use without asking the data subject for consent or informing them. This contradicts the transparency and information obligations under the Articles 12 to 14 of the GDPR and the basic principles of the GDPR that seek to provide data subjects with control over their information.
In fact, Article 5 (6) shows that, by default, data subjects will not be asked to consent to this re-use and that consent is only proposed as a "last resource" solution in case other conditions cannot be fulfilled. The data subjects, whose information is at the heart of the decision, is simply removed from the discussion. This Article must be modified to prevent the erosion of data subjects' rights and involve data subjects in the process of authorising re-use of their information.

There are also serious flaws with the measures proposed for re-use in Article 5. For example, Article 5 (3) states that "*Public sector bodies may impose an obligation to re-use only pre-processed data where such pre-processing aims to anonymize or pseudonymise personal data.*" However, the anonymisation and pseudonymisation techniques referred to here are never sufficient to guarantee that information cannot be traced back to data subjects. Multiple studies have shown that anonymisation and pseudonymisation techniques are insufficient to prevent re-identification of data subjects. As noted by Rocher, Hendrickx, & de Montjoye in their 2019 paper 'Estimating the success of re-identifications in incomplete datasets using generative models,' "even heavily sampled anonymized datasets are unlikely to satisfy the modern standards for anonymization set forth by GDPR and seriously challenge the technical and legal adequacy of the **de-identification release-and-forget model**."[3] Under no circumstances should datasets containing personal data, anonymised, or pseudonymised personal data be released to re-users as the risk is always present that anonymisation techniques will ultimately be rendered useless by further development in re-identification techniques.

Concerns also exist regarding the proposal in Article 5 (4) that public sector bodies may impose obligations "*to access and re-use the data within a secure processing environment provided and controlled by the public sector*." Even such secure processing environments have been shown to be vulnerable to sophisticated reidentification techniques and are thus unsuitable to secure the

[3] Rocher, L., Hendrickx, J.M. & de Montjoye, YA. Estimating the success of re-identifications in incomplete datasets using generative models. *Nat Commun* **10,** 3069 (2019). https://doi.org/10.1038/s41467-019-10933-3 (emphasis added)

protection of anonymised or pseudonymised personal data. For example, Diffix, a state-of-the-art query-based secure environment which claimed to be fully privacy preserving and GDPR compliant, was shown to be vulnerable to noise-exploitation attacks which allowed researchers "to infer private information about individuals in the dataset" with an accuracy of up to 97.1%.[4]

In short, anonymisation, pseudonymisation, and even the use of secure environments are never free from vulnerabilities. Today's state of the art will always be eclipsed and rendered vulnerable by developments in re-identification techniques, meaning that personal data cannot be securely shared for re-use by any of these means in a long term perspective. We therefore recommend that sensitive personal data, or data from which sensitive personal data may be inferred, not be shared for re-use according to the measures set out in Article 5.

3. Preventing the monetisation of data processes

Following the shift noted in the European Commission's approach to governing data from data protection to a data-driven economy, Article 6 of the DGA authorizes public bodies to charge a fee for the re-use of data. These provisions risk creating an incentive for public bodies to monetise the use of personal data. As authorising the re-use of personal data could become a source of income, it could reduce incentives to limit data processing or better protect information. To limit these risks, at minimum, data defined under Article 3.1.d (personal data) should not be covered by the scope of Article 6.

4. Establishing Data Protection Authorities as main supervisory authorities under the DGA

Throughout the DGA, the Commission refers to different competent bodies (Article 7) and competent authorities (Articles 12 and 20) that shall be tasked with the enforcement of the different provisions under the law. We argue that Data Protection Authorities should be tasked with the enforcement of these rules, if the DGA would continue to cover personal data in its scope. A large number of provisions on data altruism and data re-use directly overlap or interact with the GDPR, and involve the interpretation of consent requirements as well as monitoring the processing of data: all of which is clearly a responsibility of DPAs as foreseen under Article 55 of the GDPR, in particular.

It follows that Recital 33, Article 7, 12 and 20 shall be modified to refer to "competent authorities under Regulation 2018/679".

We recall that DPAs and the European Data Protection Board (EDPB) should be tasked with drawing a "data altruism consent form" described in recital 39, 41 and 42 and Article 22, and not the European Commission who should instead be consulted.

[4] Gadotti, Houssiau, Rocher, Livshits & de Montjoye.When the Signal is in the Noise: Exploiting Diffix's Sticky Noise. https://www.usenix.org/conference/usenixsecurity19/presentation/gadotti

Finally, as the DGA creates a European Data Innovation Board, which would include the EDPB, it shall be noted that the EDPB and its members shall be the responsible for tasks described under Article 27 (a) as it related to the application of data re-use measures.

Conclusion

The Commission's current proposal, and particularly its treatment of personal data, risks both undermining the existing data protection acquis and creating regulatory confusion by introducing measures that conflict with existing safeguards.

With the GDPR and the ePrivacy Directive, the European Union has established itself as a world leader in data protection. The Data Governance Act must build on that legacy instead of undermining or challenging it.

We remain at your disposal for any further information.

For more information, please contact:

Estelle Massé
Senior Policy Analyst
estelle@accessnow.org

Daniel Leufer
Europe Policy Analyst
daniel.leufer@accessnow.org