BDVA position paper

Response to the European Commission's proposal for a Data Act

May 2022

Executive Summary

The Big Data Value Association (BDVA)[1] is a European industry-driven research and innovation community on Big Data, Data Value, and Industrial AI. BDVA is also a unique platform for pre-competitive collaboration between industry and research, combining a research-oriented and experimental approach to data-driven innovation and a focus on competitiveness and adoption.

BDVA welcomes the Data Act and believes it will play a fundamental role in the development of European data spaces and in the creation of data spaces governance frameworks and infrastructures.

BDVA also welcomes the introduction of interoperability and standards as a key measure to achieve the goal of fairness in the distribution of value along data value chain, We believe Industry and researchers should play a strategic role in the life cycle of production, monitoring and evaluation of interoperability requirements and standards for data spaces. For this purpose we would recommend the Commission and ESOs to closely collaborate with communities such as BDVA, the Data Spaces Business Alliance (BDVA, FIWARE, Gaia-X and IDSA), the upcoming Data Spaces Support Centre (DSSC) and the existing Federation of Big Data Innovation Hubs (EUHubs4Data). The community underlines the importance of experimentation in data spaces with regards to data governance models and mechanisms.

BDVA stresses the importance to clarify the relation and interaction that the Data Act proposal will have with the existing EU digital policy and regulatory frameworks (incl. GDPR, Regulation Free Flow of Non-Personal Data, Open Data/PSI Directive, Trade Secrets Directive, security and competition law) and the ongoing negotiations (incl. AI Act, ePrivacy Regulation, Digital Services Act).

It is also advisable to work towards a taxonomy of data categories and clarity on how to deal with different data categories. Further clarification is also needed on the types of data concerned by the different parts of the proposed text, and a clear understanding of what type of data is targeted in the various chapters of the Act. Additionally definitions of some of the roles and actors and related concepts may be necessary in chapter I. This is particular important of the "Operators of Data Spaces".

The Data Act is expected to incentivize data sharing and build trust to promote value in the data economy. However, this will only be achieved if regulation acts as an enabler, rather than as a set of restraining obligations. If businesses see additional costs for them to provide data or risks of penalties for non-compliance, they will make less data available leading to a reduction of the possibility for value creation and for common good.

The proposed provisions aim for facilitating access to and use of IoT data by consumers and businesses, while preserving incentives to invest in ways of generating value through data.

BDVA stresses the importance of clear definitions and a coherent identification of scope of these provisions and suggests improvements in the scope and definitions of data, products, and roles, and in the data holders´ rights protection.

Concerning the switching in between cloud services, we acknowledge the intention of the Data Act to foster and safeguard a maximum level of switching including data, applications, and any other data asset for customers of a data processing service provider). However we suggest important improvements in the definitions of key terms, in the deadlines proposed, in the term and usage of the

[1] From 2021, the legal name of BDVA changed to DAIRO - Data, AI and Robotics aisbl. BDVA continues in full use as the brand.

Data, AI and Robotics (DAIRO) aisbl
Avenue des Arts, 56 ● 1000 Brussels ● Belgium
E-Mail: info@core.bdva.eu ● Web: www.bdva.eu

"functional equivalence" and we propose the extension of norms held in data sharing to applications that are transferred to data as well.

As for the link with the development of trustworthy AI, BDVA underlines the connection between data and AI and stresses the importance of further investigating the impact of the proposal on the life-cycle of data-driven AI (e.g MLOps).

Our response to the feedback request covers several aspects of the proposal for a Data Act and sees it as a needed resource but also as a piece of legislation that needs improvements to deliver a fair and competitive data economy in Europe. The feedback of BDVA to the Data Act is structured as follows:

1. General remarks
2. Better access to IoT data
3. Tackling unfairness in contractual clauses
4. Make business data available for the common good
5. Easier switching between cloud services (and edge)
6. International context non-personal data safeguards
7. Facilitate data flows through technical standards and interoperability
8. Data Act and Data Spaces
9. Data Act and Trustworthy AI

1. General remarks

Data Act proposal in the wider regulatory context

The Data Act proposal, like the Data Governance Act proposal before it, relies on 'fairness', which is being developed as a legal obligation falling upon economic actors. The operationalisation of fairness, however, raises many challenges. The Data Act proposal is at the intersection of many legal frameworks (European and national) where 'fairness' is interpreted in various ways, playing different roles.

BDVA/DAIRO therefore stresses the importance to clarify the relation and interaction of the Data Act proposal with the existing EU digital policy, regulatory frameworks (incl. GDPR, Regulation Free Flow of Data, Open Data/PSI Directive, Trade Secrets Directive, security and competition law) and ongoing negotiations (incl. AI Act, ePrivacy Regulation, Digital Services Act).

Most notably, the alignment with GDPR raises concerns. The objectives of the Data Act can challenge the compatibility with GDPR's data minimisation and purpose limitation principles. The Data Act acknowledges that the GDPR shall prevail in case of contradiction and reiterates the distinction between 'personal' and 'non-personal data'. In practice, however, this distinction can be hard to make. Whether such safeguards are sufficient to protect fundamental rights of individuals remains insufficiently assessed and the absence of clarity on the interaction can have a chilling effect on the realization of the objectives of the Data Act.

Another question relates to whether consistency and link with Artificial Intelligence (AI) and Machine Learning (ML) models should also be made. With the Data Act, federated data spaces will be enhanced to leverage AI deployment. To this end, AI and ML solutions should be linked with the Data Act and an investigation upon their limitations in the framework of the Data Act should be carried out.

Clarification is also needed on the impact of the proposal on national law, where data sharing provisions may be found in a large number of sectors.

Relatedly, it is still hard to picture how the roles identified in the Data Act ('data holder', 'data recipient', 'third party') align with roles identified in other legislations, *i.e.*, the Data Governance Act proposal (where the term 'data recipient' is not used, but the term 'data user' seems to have a similar meaning) and the GDPR.

Relation between the protected interests in the Data Act proposal

The Data Act aims to protect, i.a., consumers and small businesses, as well as public authorities when dealing with exceptional circumstances demanding data. However, the interests of these categories of actors are not necessarily aligned, so that choices are made in the Data Act to protect some (often SMEs) to the detriment of the others (potentially consumers and public authorities), see for instance Art. 7 and 14(2). However, such a choice is debatable. In particular, it appears to be too one-sided while not taking specific circumstances into account (I;e. when the provision of data by an SME to a public authority would not incur disproportionate costs and risks).

Defining the concept of Data

The Data Act aims at addressing a vast category of activities deemed essential to facilitate the sharing of data among private and public entities. By doing so, and in an attempt to widely embrace the vast diversity of situations, it tries to regulate activities that are very different in nature and in objective. For instance, portability conditions between private actors in cloud is not comparable to B2G data sharing.

Therefore, such a diversity demands a clear understanding of what type of data is targeted in the various chapters.

Different types of data sharing (business to business, business to public organization, public to public organization) require different handling what is not sufficiently addressed by the DA.

The first incentive to make data available is to create the trust needed by all operators. It demands a clarification and a clear delimitation of;

- what is shared and what is not,
- what is ported and switched, and what is not,
- what is requested by public authorities (whether in support of their basic public services or in use for research) and what is not.

Consequently, as such, we would welcome a clarification in the data taxonomy throughout all the proposal.

2. Better access to IoT Data

Definitions and scope

The proposed provisions aim for facilitating access to and use of IoT product data by users, whether consumers or businesses, while preserving incentives to invest in ways of generating value through data. BDVA stresses the importance of clear definitions and a coherent identification of scope of these provisions.

With regards to the definition of dat

Art. 3,1 lacks clarification whether all data generated during operation of a product must be made available or only data available to the data holder. Chapter II is based on the premise that the data holder (often, the IoT product manufacturer/related service provider) has factual control over such data, which however is not always the case.

The Data Act does not provide for a clear list or substantive scope of which data shall be deemed in the scope of art. 3 to 6. As a result, such scope is subjective: it depends on the willingness of the data holder to produce / collect / use data (and on the necessity of the product's operation). This may have both positive and negative consequences. On the positive side, it may help regulation be more future-proof and encompass (yet) unforeseen data. However, it may also disincentivize data production in the first place, in order for the data holder not to have to share them. It may also complicate enforcement as one (and especially small businesses and consumers, and some service provider) may not even be aware of the type of data which are generated by the use of the IoT products.

While Chapter II seems prima facie to apply to 'raw data', how these should be distinguished from more or less advanced forms of processed data (enhanced, inferred, derived) remains unclear. Attention should be given to investments made by enterprises to clean, sanitize, transform, store and make data accessible, otherwise it may limit the commercial desirability to invest in these activities.

It is advisable to work towards a taxonomy of data categories and clarity on how to deal with different data categories.

With regards to the definition of product:

The Data Act, including the Impact Assessment, does not provide for a clear explanation why tablets, PCs, smartphones and apps are entirely out of scope (and not, *i.e.*, solely data stored for certain reasons).

The Data Act does not include a definition for "competitor's product" (art. 4 & 5), which is however of the utmost importance for the economy and the principle of fair competition. In order to provide for workable direction, it is advisable that the 'competent authorities' in the Data Act establish a permanent connection with other specialised authorities, such as competition authorities and sector-specific competent authorities where appropriate.

The definition of "product" for the purpose of the Data Act as, essentially, IoT product, raises consistency issues with product legislation. It is therefore recommended to use the term 'IoT product'.

With regards to the definition of Data holder/ data recipient/third party

What concerns the concepts of Data Holder (art. 8, Ch. III), data recipient, third party, it is unclear from the Article what is the exact definition of it. In complex industrial systems, the roles of different private and public legal entities can change and get intertwined. A clear distinction should be made between the separation of the roles, obligations and rights related to that role, and legal entities.

With regards to data holder rights' protection

Chapter 2 does not sufficiently clarify how it relates to other legal frameworks related to data. In particular, it is unclear from Art. 4(4) whether trade secret protection will concretely take precedence over data sharing obligations under the Data Act and how it will interact with different types of trade secrets and, possibly, other confidentiality obligations. The Data Act is indeed entirely silent on the effect of other types of secrecy, i.e. for (cyber)security reasons.

Sharing IoT Product data puts manufacturers at risk that such data could be used by competitors. The Data Act does reckon the risk and does adopt a few safeguards (see Art. 4(4), 5(4), 6). However, it is difficult to keep track of data in practice and therefore to ensure that such provisions can be enforced. This is for this reason, it seems, that Art. 11 allows data holders to engage in technical protection measures ('TPMs'). However, first, the scope of Art. 11 should be clarified, in particular, whether it applies solely to Chapter 2 situations or to, potentially, any data contract (in which case it could lead to undue data appropriation). Second, Art. 11 should be consistent with other TPM provisions in EU law (i.e. in copyright law). Third, Art. 11 should not be viewed as an exhaustive enforcement of the rights of data holders under chapter 2. In other words, it should not be perceived, i.e., by enforcement authorities, as a reason for *not protecting* the rights of data holders when infringed.

Beyond the sui generis right on databases, it is also advisable to clarify in the Data Act whether IPRs and/or trade secrets, i.e., applicable to the environment of data, could stand in the way of Chapter 2. For instance, copyright on software including APIs could require a mandatory license for Chapter 2 to be applicable. Especially in multi-stakeholders environment, legal certainty is necessary in this respect.

As Art. 3 deems data sharing obligations to be part of the design of the product, what would be the consequences of a breach of such obligation? Lack of conformity, impact on the market readiness of the product? Given the imbrication between such data sharing obligations and other important branches of the law (product legislation, sales of goods, contract law, consumer law), further clarification also on the relationships with such legal frameworks is advisable.

3. Tackling unfairness in contractual clauses

BDVA follows and can help facilitate the discussions on the voluntary or binding nature of data transactions and provisions regarding unfair contractual clauses.

Questions arise related to the unfair clauses provisions, comparing to consumer law:

- on the use of the term 'unilaterally imposed', how this is different from 'not individually negotiated' and the related interpretation in consumer law
- on the list of unfair clauses. The general provision of Art. 13,2 leaves other cases open, creating discretion and uncertainty in the interpretation of what is considered 'unfair'. It could also bear the risk of 'clearing' the remaining clauses from consumer law which are not deemed expressly unfair under the Data Act. This could be clarified in a recital or in the body of the text.

The deemed unfair contractual clauses are, essentially, *not* data-specific, in the sense that they could be imposed with other contractual objects. This points to a question why they would be regulated from a data-targeted perspective. It comes with the risks of affecting legal certainty and consistency of challenging the future-proof capacity, however instrumental to protecting weaker parties (such as SMEs, beneficiaries of the said provisions). Further assessment in practice and with stakeholders could be proposed on the topics of horizontal rules and data-specific unfair clauses.

BDVA welcomes the plans to develop model contract terms for B2B data sharing and can propose valuable experience and contributions from the community to identify and develop the balance in contractual rights and obligations and to co-create with the community stakeholders these model contract terms, leveraging the research and deployment in the multiple pilots of BDV PPP projects, industry agreements and codes of conduct.

4. Make business data available for the common good

The Data Act is expected to incentivize data sharing and build trust to promote value in the data economy. However, this will only be achieved if regulation acts as an enabler, rather than as a set of restraining obligations. Where businesses see additional costs for them to provide data or risks of penalties when providing data, the risk of non-compliance could increase. Companies will make less data available and reduction of the possibility for common good.

We would welcome clear provisions regarding the use of such data by researchers. We would also welcome provisions and safeguards against the re-transfer of such data between research organizations with a clear right of the initial Data Holders to be informed in case of re-use of data upon transfer of the results further to third parties or even upon sharing to the general public.

It is arguably difficult to agree on a market price for certain sets of data (see as per Art. 15), as there may be cases where data is offered by only one participant. Specific incentives should be elaborated and introduced by the act to share certain datasets for reasonable price or free of charge when it is possible.

Supervision on data sharing and shared responsibilities between different actors are not sufficiently addressed by the Data Act. Data Sharing operating model and framework for compensation model(s) are not clearly introduced and defined. Such an operational model is however required to address the question ranging from who is responsible for what, and the extent of such responsibility, through who is eligible to supervise such tasks (including the decisions on trustworthiness if data is misused), tackling risk assessment and plan, roles and responsibilities for different legal entities. A forum should be established to escalate problematic situations and agree upon compensations.

5. Easier switching between cloud services (and edge)

With regard to cloud services, a self-regulatory approach was introduced in the Free-Flow of Non-Personal Data Regulation, which hasn't drastically affected markets dynamics, and hasn't provided the expected more open approach for non-personal data exchange. Therefore, another approach is being introduced with this proposed Data Act, namely an explicit regulatory one.

The Data Act should clarify that it creates a right to switching at any time when it is needed between different services (within the boundaries of the provider's responsibility).

The impact to move from the self-regulatory approach (from previous CoCS and SWIPO.eu) to the specific obligations as per the Data Act will require quite some effort from the market. For the cloud and edge service provider market, there will be challenges as they become responsible for providing support far into the switching process, possibly beyond the duration of the contract. The preparation of data for exchange, including the necessary interoperable solutions in an appropriate and secure manner, requires a lot of effort in practice. Especially SMEs and specialized industrial companies do often not have the competences and sufficient resources, even if they have the same rights to port data. It can lead to unfair competition between users, whether they are SMEs and larger enterprises.

We acknowledge the intention of the Data Act to foster and safeguard a maximum level of switching including data, applications, and any other data asset for customers of a data processing service provider. The key terms, though, are (i) either too wide or not provided at all, (ii) only obligate the original service provider (not even the customer itself) or are (iii) so restrictive (e.g., 30 days maximum notice period) that this will considerably impede the development of new and materially reduce profitability of any existing European service provider.

The hard deadlines, proposed by the Data Act, should be turned into a more general phrasing (I.e. referring to 'reasonable deadlines for transfer' depending on the nature and length of the contract, possibly agreed upon by the parties, etc.). This should future-prove the Data Act, make it more adaptable to the various types of contracts and align with Contract Law and its principle on freedom of contract.

The term "functional equivalence" (Art. 26) regarding the different infrastructural software services solutions is particularly vague. A broad range of solutions will be present also in the coming years. It makes it difficult to measure the equivalence in comparable KPIs and metrics and could prevent the Data Act from being future-proof. It is also unclear which criteria this equivalence should be based on, I.e., how functional equivalence should be interpreted where software solutions are based on completely different technological stacks.

We would welcome clarification on the safeguards at the disposal of the provider against misplaced complaints that no "minimal set of functionalities" was delivered after switching. Similarly, we would welcome clarification on the validation of the switching process, so the provider is not kept hanging.

More generally, ensuring full continuity between the end of the prior service and the new service provided by different parties, whatever the choice made by the user, is unrealistic.

While 'data sharing' is often referred to, it should be recalled that, technically, it is the application (computational workload representing a model for data analysis) that is being transferred to data instead of data being copied to the premises where data processing takes place. This implies the presence of trade secrets and/or IPRs, which should be preserved. The relationship between the switching provisions and trade secrets/IPRs should thus be further clarified.

6. International contexts non-personal data safeguards

Article 27 targets both illegitimate access to data from non-EU countries and public request for data by a non-EU country. It begs noting that a data request by a non-EU public authority could be targeted at any entity submitted to foreign domestic legislation, such as US law. This applies not only to Cloud providers but to a broad range of actors, who could therefore find themselves in a tricky situation of conflict of law. Against this background, it should be recalled that the best solution remains the negotiation and adoption of an appropriate executive agreement.

As mentioned in Chapter 1, data, as tackled by the DA, is a broader concept than only personal Data. It includes commercial/ trade secrets and classified data. It is necessary that a cloud service provider takes all possible measures to make sure that data is not illegitimately accessed outside of the EU.

Art. 27 should address topics like purpose limitation, data minimization etc, especially when links to sensitive data and persons (I.e. in the context of Law Enforcement) are to be expected.

Enforcement measures are currently not included in the scope of the Data Act. It remains therefore unclear how to enforce the international data sharing ban. Opposing a judicial or and administrative provision from their country of origin or a third party could be difficult for providers of data-related services. This challenge could be tamed by the possibility, set in paragraph 3, to rely on public authorities to evaluate the risk in any data request made by authorities outside the EU and to help define the minimal scope of data to be communicated if authorized. Indeed many member States do already have domestic legislation prohibiting communication of some data to foreign countries (such as the 1968 Blocking Law in France). Ultimately, such public authorities should, whether based on an executive agreement on data access or not, act systematically as proxies for providers receiving such requests. Another possibility would be for the legislator to provide a compliance mechanism, other than - and further-reaching than - a mere advice on the transfer. For example, a provision may be inserted to give foreign companies a stronger leverage to oppose their government, such as a judicial decision from a European Court. In the case of the GDPR, the enforcement mechanism is just limited to the stopping of all transfer to third countries. It is unknown, however, if this solution will be adopted in the case of the Data Act, and how it will be enforced in practice.

The general positioning of the Data Act with respect to international regulations as well as the impact of the Data Act on the associated partners of the EU is not clear yet. A link should also be clearly made with the other EU regulations dealing with international transfer of data, I.e. Art 30 of the Data Governance Act, art 48 GDPR. With all new such regulations having different scopes and apparatus, it is not easy for providers to navigate the rules. The position of the Data Act towards potential global cloud-based platforms with open data should be clarified.

Not all countries outside of EU with such intrusive legislation have agreements with EU countries on accessing data from the EU. The EU should try to ensure that the same terms of contract will be applied also to the international operators and actors. A clear separation of concerns between the US Cloud Act and Data Act should be introduced, if only to take into account the fact that many non EU-Countries have no similar legislation. The estimation of risks concerning the US Cloud Act requires further discussion at EU level. In particular, there is a need for an EU-US agreement on law enforcement data access, as big service providers are US based, and regulations on the validation of extraterritorial effect of US law.

It is also not clear in the Data Act how to regulate the application of law concerning the requested provision of data. For example, data can be requested to be provided by the court decision from the third country. The mechanism of provision and the limitations on provisioning are not tackled in the Data Act yet.

While performing the transfer of data, users and data holders (in case data provider is another organization than data holder) should have a right to be informed of the sharing supported by the norms of DA. Public bodies are prescribed to be transparent as well when requesting data in cases of exceptional need, clearly stating the purpose of the request, the intended use of the data and the duration of that use. However, in such cases no obligation of informing the Users is laid down by the current version of the proposal, unless the requested data set contains personal data, and this circumstance seems at odds with the general transparency obligation imposed in the rest of the act.

7. Facilitate data flows through technical standards and interoperability

Definitions

Interoperability provisions of the Data Act (chapter VIII) apply to data spaces (art. 28), data processing services (art. 29), and smart contracts (art. 30). The Act describes specific duties, roles

and responsibilities for various actors. Definitions of these actors and related concepts may be necessary in chapter I, This applies in particular and most importantly to the definition of “Operator of Data Spaces” and the concept “Data Space”.The same is needed for the notions that describe the type or interoperability requirements. Art. 29,2, mentions a list of notions that are not defined at all and not always set in precedence (such as existing Cloud ISO standards).For instance, the notion of “Cloud application aspects of application syntactic portability”, or “application instruction portability” Definitions will help to understand scope and facilitate the work mandated to the European Standard Organisations.

Standards development, enforcement and appropriateness of standards

With Chapter VIII, EC can task European Standardisation Organisations (ESOs) to draft harmonised standards or adopt common specifications in cases where harmonised standards don´t exist or are not sufficient. The Commission can also adopt guidelines to lay down detailed interoperability specifications (e.g architectural models and technical standards). All this allows the EC to mandate interoperability.

The provisions in art. 28 contain information on the regulatory framework or architecture of the European common data spaces, and mandates that essential requirements are adopted by Data Spaces Operators (dataset content, use restrictions, data structures, and technical means accessing the data). We believe Industry and Research players should be strategically involved in the life cycle of production, monitoring and evalution of interoperability requirements and standards for data spaces (extended to data processing services and smart contracts). For this purpose our recommendation is that the EC works closely to communities such as BDVA, the Data Spaces Business Alliance (BDVA, FIWARE, Gaia-X and IDSA), the upcoming Data Spaces Support Centre (DSSC) and the existing Federation of Big Data Innovation Hubs (www.euhubs4data.eu)

The development of common specifications and standards, as outlined in recitals 79 and 86, can take a lot of time. In order to create an outlook for stakeholders involved, it would be beneficial to have a tentative timeline for the standards development. Additionally, we believe It can be expected that time for compliance with interoperabitliy requirements will be longer than a year (referring torecital 89).

Finally, the text consistently addresses ‘European standards’. To avoid fragmentation, an inclusion of international standards would be very relevant to address as well.

8. Data Act and Data Spaces

The Data Act and Data Spaces are important pieces of the puzzle to realise the ambition of the European Data Strategy and the vision for a European internal market for data. BDVA fully supports these objectives and helps facilitate the development of common data spaces, collectively creating an interoperable data sharing environment, enabling data sharing and (re)use within and across sectors in a secured, trusted and fair way, fully respecting EU values and norms.

The community recognises the importance of fairness for the design of data spaces and therefore acknowledges the objectives of the Data Act to ensure fairness in the allocation of value from data among actors in the data economy and to foster access to and use of data. The Data Act will contribute to the creation of the data space governance frameworks and infrastructure.

BDVA underlines the importance to contribute to the empowerment of data users and data holders with respect to data and to establish a balance between the rights and interests of all stakeholders involved, with a general objective to make wider use of data possible for a broad range of actors.

The community would welcome clarifications on the concepts of data space and data space operator in the Data Act proposal (on this, see also 7).

The community will continue to engage its wide network of members in a multidisciplinary and collaborative approach to identify challenges and provide recommendations on the impact and future operationalisation of the proposed act and its objectives in the context of data spaces.

The need for a consistent and agile legal infrastructure for data spaces is stressed. As indicated in the sections above, challenges arise from the complexity in legal frameworks and their interaction and the complex and competitive relationships between the stakeholders within data spaces. This underlines the need to complement technical interoperability with legal interoperability. The awareness of the need for scalable contracts is welcomed. Finally the community underlines the importance of experimentation in data spaces from different sectors, with different purposes with regards to data governance models and mechanisms.

9. Data Act and Trustworthy AI

The increasing availability and greater importance given to data as the main training and testing resource for Artificial Intelligence algorithms has resulted in an avalanche of new technological applications that are fed with data.

Access to data is needed for the whole lifecycle of Artificial Intelligence algorithms, however, the purposes for the use of such data differ and trustworthiness of data can have different meanings, depending on the stage of lifecycle of AI application (development of the model, training and validation of the model, and, the deployment and running the model in operational settings). Futher investigations on the impact of Data Act on all these stages is needed. Particularly relevant is the third stage (deployment and running model in operational settings) whereaccess to data is constantly needed for the maintenance of a trustworthy AI model in operational conditions,

Good quality data is essential for all stages of the lifecycle of AI model. For Trustworthy AI this is even more relevant because the requirements for trustworthy AI are much higher and strictly linked with data quality. Specifically, the following key requirements, from the AI High Level Expert group on Trustworthy AI, are relevant for the Data Act:

- Privacy and data governance: it should be considered that non-personal data can sometimes be related to personal data (for instance geospatial or location data).
- Transparency: the provenance and collection process of data that is transferred from platforms and reused should be described in metadata.
- Diversity, non-discrimination and fairness: it should be possible to know how data from platforms is collected so that it can be assessed whether the data is unbiased
- Accountability: any damages coming from previously collected data that is transferred under the data act should be traceable to the collection and processing.

These requirements lead to the following considerations:

- Low-quality data or data without appropriate background information can be misinterpreted and this can lead to damages. In order to ensure a trustworthy AI system, a minimum level of metadata (meaning, accuracy, provenance, quality) should be provided.
- It should be considered that non-personal data can sometimes be related to personal data (for instance geospatial or location data). It should be explicitly mentioned that the transfer of non-personal data may create infringements on the GDPR.
- The provenance and collection process of data that is transferred from platforms and then further reused, should be described in metadata. This is necessary to ascertain that the diversity, non-discrimination and fairness requirement is met. Even when the data itself is non-personal, it may become personal upon further combination. In such case, it could lead to unwanted results - and therefore untrustworthy AI.

Fulfilment with the above should also support accountability in case of damages stemming from the usage of previously collected data that is transferred under the Data Act.

Clarifications brought to the terms used in the Data Act and notably to the definitions of Data Holder, Data recipient and Third party, could both provide transparency concerning data and enhance an easier access for the user. Transparency is an important principle to ensure that the compensation requested by the data holder is reasonable, or, in case the data recipient is an SME, that the compensation does not exceed the costs directly related to making the data available.7

About this Document

Main editors (in alphabetical/surname):

- Natalie Bertels (imec/Ku-Leuven, Researcher, BDVA TF5 co-lead)
- Freek Bomhof (TNO, Business Consultant and Project Manager, BDVA TF5 co-lead)
- Elena Lazovik (TNO, Scientist)
- Francesca Manni (Philips, AI Scientist, TF7.Healthcare lead)
- Richard Stevens (IDC, Director)
- Amal Taleb (SAP, EU Public Affairs Director)

and from the BDVA office (alphabetical order)

- Ana García Robles (Secretary General)
- Mattia Trino (Operations Manager)

This paper is the result of a cooperative work that gather inputs from BDVA/DAIRO members that participated in the following meetings and workshops with over 50 orgnisations involved:

- BDVA Activity Group 49 (16/03/22) - Session by BDVA Task Force 5.Policy&Societal: The new policy developments at EU level
- BDVA Poistion Paper Data Act workshop (1st general workshop on 23/03/2022)
- BDVA Poistion Paper Data Act workshop (Second workshop on 28/04/22)

Additionally BDVA members have contributed with input through an internal consultation (run in between the 1st and 2nd workshops listed above). Many thanks for the special contributions from City of Espoo, Huawei, imec-KULeuven, IDC, NTT Data, Philips, SAP, Software AG, TNO, TU Eindhoven, VTT and all BDVA/DAIRO members that have contributed with comments and suggestions to this endeavour.

About BDVA

The Big Data Value Association – BDVA, (from 2021, DAIRO - Data, AI and Robotics aisbl), is an industry-driven international not–for-profit organisation with more than 230 members all over Europe and a well-balanced composition of large, small, and medium-sized industries as well as research and user organizations. BDVAfocuses on enabling the digital transformation of the economy and society through Data and Artificial Intelligence by advancing in areas such as big data and AI technologies and services, data platforms and data spaces, Industrial AI, data-driven value creation, standardisation, and skills. BDVA has been the private side of the H2020 partnership Big Data Value PPP, it is a private member of the EuroHPC JU, itis also one of the founding members of the AI, Data and Robotics Partnership and a partners in the Data Spaces Business Alliance. BDVA is an open and inclusive community and is always eager to accept new members who share these ambitious objectives.

Contact for further information: info@core.bdva.eu