

TOWARDS A EUROPEAN-GOVERNED DATA SHARING SPACE

Enabling data exchange and unlocking AI potential

BDVA Position Paper v2

November 2020



BDV BIG DATA VALUE
ASSOCIATION

EXECUTIVE SUMMARY

The BDVA's vision of a European-governed data sharing space can mark a milestone in the evolution of the European data economy. Rather than referring to data as the new oil, we think of it as a new currency underpinning a new data economy, catalysing the entire European economy to develop faster and more effectively. Opportunities presented by Artificial Intelligence (AI) can now be considered as the primary driver of the data economy. These opportunities cannot be fully exploited without wider access to large volumes of high-quality data; which remains severely limited. We consider widespread, secure and effective data sharing as the only solution that can realistically address this bottleneck.

The envisioned European data sharing space, first introduced in the earlier version of this paper¹, refers to a space that is composed of, or connects, a multitude of distinct spaces that cut across sectoral, organisational and geographical boundaries. We have revisited that vision to emphasize the unique opportunity that Europe has to govern and consolidate this space at an international level.

Europe's two major internationally distinguishing realities in this area, i.e., a diversified business landscape² and an ethics-driven regulatory framework, should be considered as strengths, rather than weaknesses. If appropriately tapped into, Europe's higher business diversity can enable it to compete with regions dominated by large players that monopolise the market and possess large data assets. Furthermore, rather than be seen as restrictive, legislation enforcing European ethical values should be considered as a unique mid-term competitive advantage in the global data marketplace, provided that there is continued investment in the related R&D. Thus, values like democracy, privacy safeguards and equal opportunities can become the trademark of European data economy technologies, products and practices. European-governed data sharing must therefore reflect and respect European ethical values and principles; including openness, democracy, privacy protection, fair and equal opportunities. The strategic and timely steering of these activities towards an achievable outcome can position Europe as a global pioneer of data sharing innovation.

The Big data Value Association (BDVA), private counterpart of the European Commission in the Big Data Value Public-Private-Partnership (BDV PPP) and one of the main promoters of the Horizon Europe AI, Data and Robotics Partnership candidate³, and the BDVA members, have a strong role to play in steering and guiding these activities at European level. In full alignment with the first release this paper intends to support the dialog among European and national policy makers, industry, research, public sector and civic society in the definition of a common roadmap for the development and adoption of a pan-European Data Sharing Space, and keep guiding public and private investments in this area in particular in the context of Horizon Europe (HE)⁴ and Digital Europe Programme (DEP)⁵. The content of this paper is fully aligned with the third released of the Strategic Research, Innovation and Deployment Agenda (SRIDA) for the European AI, Data and Robotics Partnership⁶ and outcomes of the ongoing multi-stakeholder dialogue will be reflected in future versions.

To speed up the realization of the envisaged data sharing space, the BDVA considers the joint intervention by all the above-mentioned stakeholders a necessity. The envisaged macro-level space can incorporate multiple existing vertical, cross-sectoral, personal and industrial data spaces, offering services and experimentation opportunities to all stakeholders. In order to achieve this objective and generate the

1 https://www.bdva.eu/sites/default/files/BDVA%20DataSharingSpace%20PositionPaper_April2019_V1.pdf

2 https://ec.europa.eu/eurostat/statistics-explained/index.php/Statistics_on_small_and_medium-sized_enterprises

3 https://ec.europa.eu/info/horizon-europe/european-partnerships-horizon-europe/candidates-digital-industry-and-space_en

4 https://ec.europa.eu/info/horizon-europe_en

5 <https://ec.europa.eu/digital-single-market/en/europe-investing-digital-digital-europe-programme>

6 <https://ai-data-robotics-partnership.eu/wp-content/uploads/2020/09/AI-Data-Robotics-Partnership-SRIDA-V3.0-1.pdf>

broadest possible impact, concrete actions need to be urgently taken. The main bottlenecks to this vision, as observed through multiple forms of consultation with key stakeholders, remain of a business, legal and organisational, rather than technical, nature. Specific actions are needed to raise the trustworthiness, reliability, safety and visibility of data sharing spaces offering cross-sectoral, unbiased, high-quality and trustworthy data. At the same time, Europe's socio-economic realities (diversified business landscape, ethics-driven regulatory framework, etc.) presents the European establishment a unique opportunity to drive convergence efforts between several independent and international initiatives that understand the importance of realising next-generation *data sharing spaces*⁷. Convergence is required in both regulatory, research and innovation policies as well as technology and data ecosystems with the intent of developing implementable strategies, frameworks and technical solutions that can sustain an ethical, open and democratic European data economy. The implementation of advanced and rigorous European solutions can be extended internationally to realise the BDVA's vision and position Europe at the forefront of data and AI innovation.

After revisiting the identified opportunities introduced by the described vision and the challenges that stand in the way, we include a concrete call to action targeted at strategic stakeholders that can jointly guide and drive the successful development, implementation and widespread adoption of a European-governed Data Sharing Space.

⁷ In the previous version we disambiguated between different complementary technology categories behind emerging data sharing ecosystems: data platforms, spaces and marketplaces. Since then the term Data Sharing Spaces has grown into the primary catch-all key phrase encompassing all relevant concepts and technologies.

Table of Contents

Executive Summary	2
Vision	5
Opportunities in the next 10 years	7
Foreseen Challenges	8
Annex 1. Identified Challenges (updated since version I)	12
<i>Technical Challenges</i>	12
<i>Business and Organisational Challenges</i>	13
<i>Legal Compliance Challenges</i>	14
<i>National and Regional Challenges</i>	15
About BDVA	16

VISION

The realisation of a functioning and frictionless European-governed data sharing space⁸ that can successfully generate economic **value** by broadening data access for AI, relies on carefully planned iterative implementation strategies and a timely concerted effort between all relevant stakeholders. As depicted in Figure 1, the success of widespread data sharing activities revolves around the central key concept of **trust**: in the validity of the data itself and the algorithms operating on it; in the entities governing the data space; its enabling technologies; as well as in and amongst its wide variety of users (organisations and private individuals as data producers, consumers or intermediaries). To achieve the required levels of trust, each of the following five pillars must meet some of the necessary conditions:

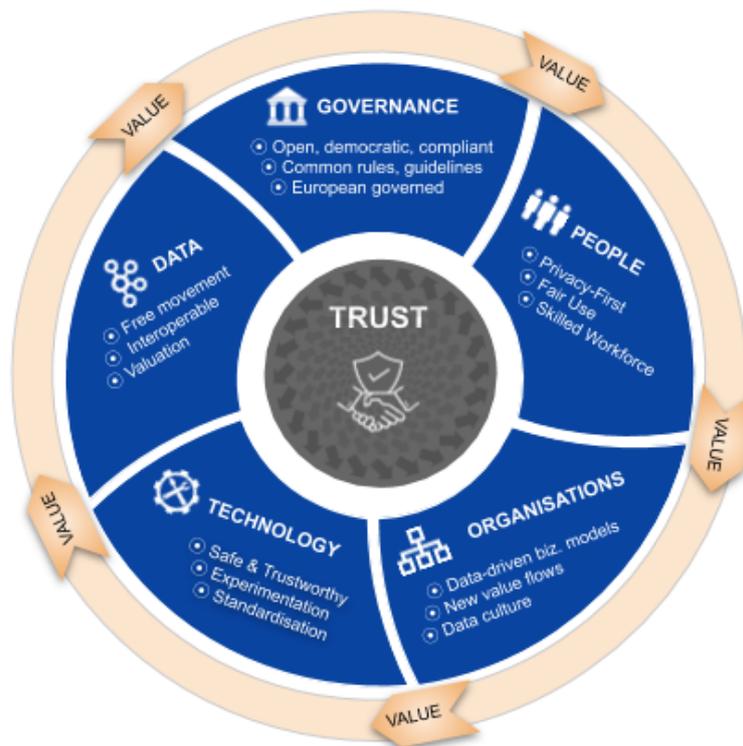


Figure 1 The Data Sharing Value ‘Wheel’ - core pillars and principles of the envisioned European-governed Data Sharing Space that generate value for all sectors of society.

1. **Data** - As a touted 5th European fundamental freedom⁹, free movement of data relies on organisational data strategies that embed methodologies for data sharing by-design (e.g., interoperability) and clear standard guidelines that help determine market value of data assets.
2. **Governance** - A European-governed data sharing space can inspire trust by adhering to the more advanced European rules, guidelines and regulations and promote European values. Participation should be equally open to all and subject to clear and fair rules of conduct.

⁸ As a singular but federated virtual space connecting several other interoperable spaces.

⁹ As proposed by ,e.g., EU member state Estonia: <https://e-estonia.com/free-movement-of-data-as-the-5th-fundamental-freedom-of-the-european-union/>

3. **People** -Data sharing needs to guarantee individual privacy and offer fair value or compensation of shared personal data. In order for Europe to drive data sharing activities, the European workforce needs appropriate reskilling and upskilling to meet the evolving labour market’s needs.
4. **Organisations** - More organisations (including business, research and governmental) need to rethink their strategy to fully embrace a data culture that places data at the centre of their value proposition, exploring new data-driven business models and exploiting new data value flows.
5. **Technology** - Safer experimentation environments are needed to catalyse the maturation of relevant technology behind trustworthy data, data access and algorithms (privacy, interoperability, security, and quality). Standardisation activities need to adjust for faster reaction times to emerging standards and the identification of new ones.

The BDVA recognizes two complementary high-impact opportunities that can materialise as a result of timely interventions to converge data sharing initiatives in Europe and realise its vision:

1. Achieve **wider access to data** so as to realise the full potential of emerging AI technology through the design and implementation of a common, trustworthy, decentralised data space that enables safe and democratic data sharing and boost the European data economy.
2. Achieve a **European-governed data space**, giving Europe the possibility to assume a prominent position steering international efforts to develop data and AI solutions that reflect and respect European ethical values including democracy, privacy protection, and equality¹⁰.

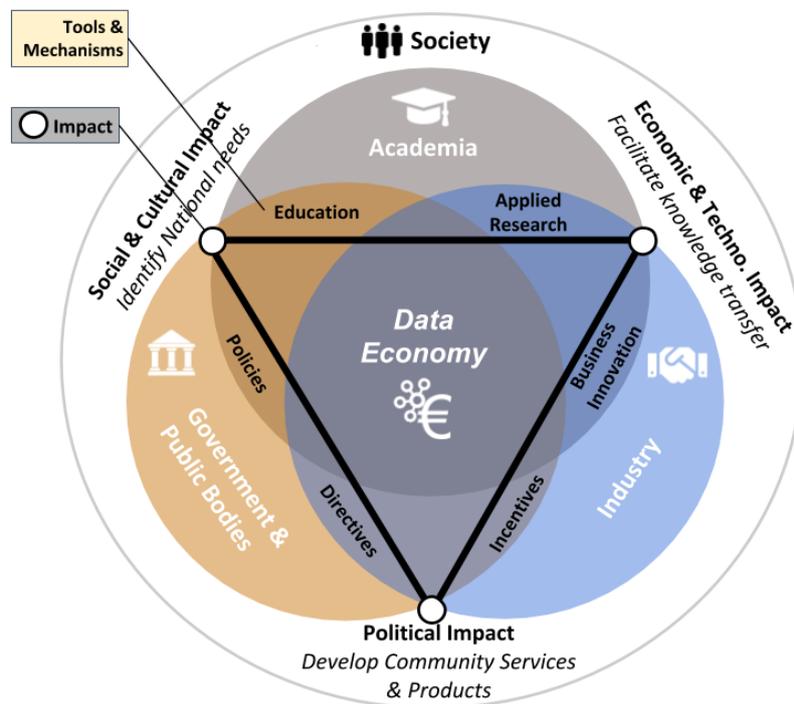


Figure 2 Tools and mechanisms that strategic stakeholders can make use of to jointly realise a data sharing space, boost the European data economy and create various lasting societal impacts

¹⁰ <https://www.euractiv.com/wp-content/uploads/sites/2/2018/12/AIHLEGDraftAIEthicsGuidelinespdf.pdf>

The introduced vision for a European-governed data sharing space needs to be built around and in consultation with the same wide array of stakeholders that can exploit its benefits as users. Figure 2, designed on the triple helix view of research, development and innovation production, shows the various roles that the active strategic stakeholders (Industry, Academia, Government) can play in the realisation of this vision, through which tools they can actively contribute and the existing potential to achieve different kinds of societal impact (Economic, Technological, Political and Cultural).

Rather than focusing on specific Business to Business (B2B) scenarios, or restricting the vision to specific sectors, we envision a data sharing space that is open to all, thus offering equal opportunities and spanning all societal spheres, including private citizens. Even though the latter are not actors in the realisation of the data sharing space, they still play an important role in data sharing. Although as the main economical driver we retain business at the centre of our recommendations, in addition to B2B cases we also consider Business to Governance and vice versa (B2G, G2B), Business to Science and vice versa (B2S, S2B) as well as Consumer to Business (C2B) opportunities.

Following the feedback received and the various consultations¹¹ that took place since the publication of the first version of this position paper, the original recommendations were translated into concrete actions that can be feasibly materialised over the next decade alongside the Digital Europe and Horizon Europe Programmes (2021-2027). The BDVA's call for action is entirely aligned with the European Commission's latest Data Strategy¹². The recommended actions take into consideration the previously identified societal opportunities and the challenges that can limit their impact, both of which are considered below.

OPPORTUNITIES IN THE NEXT 10 YEARS

In the first version of this position paper, we extrapolated trends to outline opportunities arising over the next decade for the four primary societal spheres of: industry (Business); private users (Citizens as consumers); research and academia (Science) and local, national and European government and public bodies (Government). The identified opportunities have a wide socio-economic value: possibilities to boost the data economy will lead to an improved general economy thus benefiting society as a whole. In this revised version we include a summary of the identified opportunities. Readers are kindly referred to the previous version for more details.

Opportunities for Business

- Open data marketplaces that level the playing field for industrial data sharing.
- Increased availability of vast and heterogeneous data ecosystems for AI.
- Innovative data-driven business models enabled by new value ecosystems.
- Opportunities to tap into 'safe' personal data.

Opportunities for Government and Public Bodies

- Data commons for better government services.
- AI-enhanced digital services.
- Real-time European statistics.
- Lean business environment enabled by access to government services.

¹¹ Refer to the List of Milestones summary in the About BDVA section

¹² <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1593073685620&uri=CELEX%3A52020DC0066>

- Evidence-based policy making.
- Data as evidence of Policy compliance.

Opportunities for Science

- Increasing socio-economic impact of research data across domains and borders.
- Advancing science and open innovation through data availability.
- Monetisation opportunities brought about by emerging data-driven business models.

Opportunities for Citizens

- Full control over personal data.
- Well-being and Quality of Life benefits from personal data sharing in key sectors.
- Access to personalised and cross-sectoral B2C services.
- Increased opportunities of personal data monetisation.
- New professional opportunities.

FORESEEN CHALLENGES

In the first version of this position paper we included an overview of the most critical challenges that stand in the way of the expected value generated by the identified opportunities. These challenges have been consolidated following the feedback received, and are detailed in Annex 1. The listed challenges can be categorised in two main kinds of concerns: inter-organisational (lack of suitable data sharing ecosystems) and intra-organisational (issues faced by data producers and consumers, as data sharing participants).

The most pressing inter-organisational concern remains the lack of functional and trustworthy data sharing ecosystems that inspire immediate large-scale participation. Primary causes include the lack of robust legal and ethical frameworks, as well as governance models and trusted intermediaries that guarantee data quality, reliability, and its fair use. This is compounded by the lack of widespread adherence to emerging best practices and standards (e.g., interoperability, provenance and quality assurance standards), whose maturity pace also continues to fail expectations. From a technical point of view, data sharing solutions need to better address European concerns like ethics-by-design for democratic AI, and the rapid shift towards decentralized mixed-mode data sharing and processing architectures also poses significant scalability challenges.

In terms of intra-organisational concerns, a first major concern is the difficulty to determine the value of data, due to a lack of data valuation standards and assessment tools, compounded by the highly subjective and party-dependent nature of data value and the lack of data sharing foresight exhibited by a majority of producers. The second concern revolves around the difficulty faced by data producers balancing their data's perceived value (after sharing) against risks exposed (upon its sharing) despite adhering to standard guidelines. Specific examples include the perceived loss of control over data (due to the fluid nature of data ownership, which remains hard if not impossible to legally define), the loss of trade secrets due to unintentional exposure or malicious reverse-engineering (in a business landscape that is already very competitive), and the risk of navigating around legal constraint in view of potential data policies breaches (including GDPR and exposure of private identities).

BDVA RECOMMENDATIONS: CALL TO ACTION

The first version of this position paper identified five recommended preconditions¹ for the successful development, implementation and adoption of a European Data Sharing Space. Following widespread consultation with all involved stakeholders, those recommendations have since been expanded and translated into twelve concrete actions. These can effectively be implemented alongside the Horizon Europe and Digital Europe programmes¹³. The recommended actions are categorised under five independent goals: Convergence, Experimentation, Standardisation, Deployment, and Awareness; each of which is respectively targeted towards specific stakeholders in the wider data sharing ecosystem. An effective implementation of those five goals should take place within the timeframe shown in Figure 3. Assuming the convergence initiatives that are required over the next three years¹⁴ will yield satisfactory outcomes, deployment efforts can be scaled-up with experimentation acting as a further catalyst. Other deployment efforts need to go hand in hand with intensified standardisation activities, which are key to a successful European-governed data sharing space. Activities targeted at greater awareness for all end-users can initially target organisations, entities and individuals that can act as data providers, and then extend to all potential consumers as solid progress is achieved. The actions are targeted to specific actors which map to one or more of the strategic stakeholders in Figure 2.

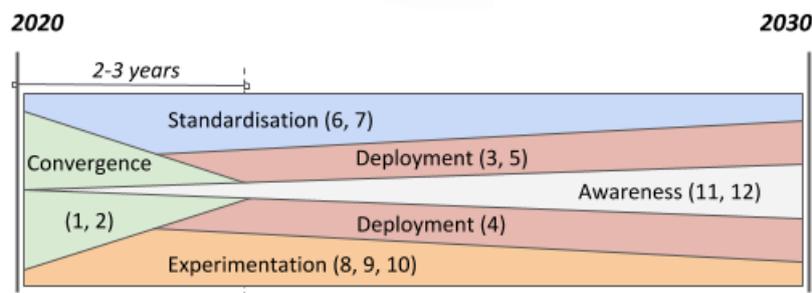


Figure 3 A suggested timeframe for implementing the recommended actions over the next decade (x-axis). The y-axis illustrates the relative, rather than absolute, effort distribution between the five phases. The absolute effort value is expected to fluctuate, reaching a peak around the 3-5 year mark.

To catalyse the **convergence** of existing national and regional concepts, efforts, priorities and strategies:

1. Launch Coordination actions to map existing initiatives at EU, member state, regional and municipal level (leveraging existing EC CSAs), and identify the common building blocks to create an impact and foster standardisation actions (see action no. 6). Efforts should particularly seek to collect, publish and systematically analyse use-cases (see action no. 8) and align legal (including tax treatments) and governance structures necessary for safe and fair level European-wide data sharing (see action no. 5).

¹³ Some actions have also found their way to various sections of the recent Strategic Research Innovation and Deployment Agenda for the AI, Data and Robotics Partnership released by BDVA, CLAIRE, ELLIS, EurAI and euRobotics, including the innovation ecosystem enablers, European AI Framework and Cross-technology Enablers sections. <https://ai-data-robotics-partnership.eu/wp-content/uploads/2020/09/AI-Data-Robotics-Partnership-SRIDA-V3.0.pdf>

¹⁴ Results of the BDVA Data Sharing Survey (August, 2020) confirm the need to act swiftly, with 39% and 52% of respondents indicating under 2 years or under 5 years respectively as a feasible time-frame for the setup and realisation of a functional European-governed data sharing space. More results here: <https://bdva.eu/task-force-1>

2. Further invest in a Euro-wide Skills Strategy to identify major skill and expertise gaps in the European workforce, devise Euro-wide reskilling and upskilling roadmaps, advise education and training bodies with remedial actions and offer further scholarships and fellowships in crucial areas of study¹⁵.

Given the nature of the above actions, they are intended for all stakeholders, but particularly the EC (and the CSA instrument), consortia behind existing EC CSAs, government at member state, regional and municipal level, industrial and research associations, standardisation bodies, consumer organisations as well as (in the case of action no. 2) educators and industry as employers.

Assuming an acceptable degree of convergence is achieved, the **deployment** of successful trusted data sharing framework and governance structures can be achieved by:

3. Funding the development of rules and guidelines for iterative, multi-phase design, creation, scale-out or merging of existing data sharing spaces.
4. Launching Research and Innovation Actions (including EC R/IAs) to evolve high-impact technology (privacy-, interoperability-, security-, quality- and ethics-by-design) with an emphasis on an interoperability initiative across existing spaces and the deployment of trust solutions for data sharing assurance.
5. Developing EU-wide Data Governance practices to clearly define rules of conduct for the fair use of exchanged data after testing different approaches to assess the impact of data sovereignty, campaign for the revision of relevant legislation, and explore means for Conformity Assessment (e.g. voluntary or licensed certification, data sharing assurance agreements) that guarantee trust.

The above actions can be primarily implemented by the EC (RIA instruments), governmental organisations, regulators, industry associations, direct data suppliers and end-users, technical and sector-oriented R&D associations, as well as (action no. 5) standardisation bodies, assurance stakeholders, and consumer/end-user organisations.

The widespread adoption of deployed accountable, reliable and trustworthy data spaces will only succeed if organisations at the forefront of relevant research, development and innovation activities bolster **standardisation** activities by:

6. Proactive coordinated engagement¹⁶ with international standardisation bodies to establish Key Performance Indicators (KPIs) that drive standardisation of common building blocks (see action no.1), successful technology (see action no.4), methods (e.g., data rights management as ownership solution) and best practices (e.g., guidelines for international data valuation, privacy preservation and FAIR principles for non-research data) behind trust-enabling solutions.
7. Assurance of sufficient conditions (assurance, market demand, government procurement) to enable widespread adherence to established standards, rules and guidelines and promotion amongst networks.

The above actions are primarily intended for technology and infrastructure providers (e.g., BDVA i-Spaces), standardisation bodies in the Information and Communications Technology (ICT) domain, sector-oriented (vertical) associations, data marketplace operators, direct data suppliers and consumers.

¹⁵ These efforts will align to and leverage results of the European Skills and Jobs Coalition and related initiatives at EU, national and regional levels. <https://ec.europa.eu/digital-single-market/en/digital-skills-jobs-coalition>

¹⁶ Using instruments like StandICT.eu.

Deployment activities should be supported and guided to exploit various **experimentation** instruments, so as to make maximum use of safe environments for testing evolutionary steps of a European data sharing space, by:

8. Investing in piloting to test-drive identified data sharing use-cases in safe and dynamic regional- and European-wide cross-sectoral scenarios (pairing member states with different levels of progress in data sharing infrastructures).
9. Engage with major initiatives offering experimentation activities that rely on data sharing (EDIH future network, BDVA i-Spaces network, etc.) to jointly explore market capabilities for sustainable new business, cross-sectoral opportunities and organizational models (e.g., for data valuation, and/or organizational transformation) and create innovative solutions.
10. Set up European regulatory sandboxes for trusted data sharing and safe exploration of risks, so as to encourage innovation and motivate data sharing.

The joint realisation of the above actions requires the attention of funding bodies within member states, industrial targets across Europe, start-ups, entrepreneurs and technological providers, direct data suppliers, consumers as well as investors, venture capitalists, incubators, in coordination with governmental organisations and regulators.

In parallel to all above actions, greater **awareness** of the opportunities offered by an open, fair and ethical data economy needs to be achieved. To this end, we call for the following supplementary actions:

11. Launch a campaign targeted at influencing organisations to revisit their data strategy and lifecycles to ensure the production of data that is sharing-ready by design.
12. Launch a EU-wide citizen-oriented campaign for an open, democratic and fair data economy and the right to the free flow of data that is both safe and trusted.

Drivers of activities for greater awareness include the EC, government at member state, regional and municipal level, sector-based industrial associations, entrepreneurs and technology providers and consumer/end-user organisations.

ANNEX 1. IDENTIFIED CHALLENGES (UPDATED SINCE VERSION I)

Technical Challenges

The ambition to realise a cross-border, cross-sectoral sharing data space and enable platforms to process 'mixed' proprietary, personal and open public data introduces new technical challenges and compounds existing ones. The impact of known challenges (e.g., the 'Vs of Big Data': Volume, Velocity, Variety, Veracity, etc.) along the data life-cycle needs revisiting following the arising opportunities for data sharing which, in addition to conventional raw data and its transformations along the processing chain, also extend to metadata, models and processing algorithms. Main challenges are:

- TC1. **Sharing by design:** *Data life-cycle management that is not designed around sharing.* The majority of data producers do not yet consider data sharing as a possibility at data creation stage. Existing data life-cycle management models need to improve how they incorporate all relevant processes, including preparing data for sharing, and finding the right data. The maturity of data services (e.g. cleaning, aggregation) in data sharing ecosystems is as crucial as the availability of the data itself; without them the data economy will not establish. Furthermore, the above differentiation between multiple kinds of data that can be made available for sharing also raises the complexity of the 'variety' challenge, and interoperability solutions need to address this change.
- TC2. **Digital Sovereignty:** *Enforcing data usage rights.* The realisation of a mixed data sharing space will only materialise if data producers are guaranteed to retain their rights as the original owners (refer to related legal compliance challenge, LC2), enabling them to retain control of who can use their data, for what purpose and under which terms and conditions. To guarantee digital sovereignty, different ownership models or suitable data rights management frameworks need to be further explored.
- TC3. **Decentralisation:** *Decentralised data sharing and processing architectures.* The need to guarantee that data producers remain in control of their data results in setups that forego data porting in favour of decentralised data storage architectures. Thus, discussions on data volumes and data velocity (data streams) need to increasingly consider both the scalability of real-time operations over dispersed data-at-rest in undetermined geographical distributions, and the distributed processing of data-in-motion which requires no intermediate storage. Standard data exchange protocols in decentralised architectures are therefore increasingly required.
- TC4. **Veracity:** *Weak verification and provenance support.* Data veracity remains crucial for the sustainability of data sharing ecosystems. Data in various stages processing chain will need to carry traceable information about its origins and operations (i.e., metadata about its raw form, algorithms and operations it was subjected to). Support for advanced provenance is required to raise trust.
- TC5. **Security:** *Secure data access and restrictions.* To enable a trusted network within which closed (proprietary, personal) data can be unlocked for exchange and sharing, issues like confidentiality and digital rights management need to be addressed appropriately. Secure access control needs to be guaranteed even within a peer-to-peer decentralised network, and therefore security solutions and exchange protocols need to be standardised across all data sharing space nodes and participants.

- TC6. **Privacy Protection:** *Maturity of privacy-preserving technologies for big data.* Although technical solutions for secure and trustworthy data sharing (e.g., Privacy-Enhancing and Privacy-Preserving Technologies, including Digital Identity management) are in continuous development, continued investment is required towards their further maturity so as to increase their reliability and uptake. More flexible ways of allowing uptake of compliance solutions also need to be explored.

Business and Organisational Challenges

Challenges in this category have predominantly been observed in industrial applications (non-personal data sharing). The clarity of the GDPR means that Industrial Data Platforms (IDPs) can potentially be extended to handle depersonalised data for added business value. The matter of legal compliance constitutes a challenge in itself, and is covered in the next section. Thus, foreseen business challenges are related to the socio-economic sustainability of a pan-EU IDP connecting multiple data spaces and offering data marketplaces, and include the following:

- BC1. **EU Values:** *Difficulty establishing EU IDPs in the global market.* EU-designed IDPs need to adhere to values such as democracy, open competition and egalitarian treatment. These characteristics can distinguish them in the global market, but also eliminate questionable ‘shortcuts’, to the advantage of global competitors. In this setting, new business models need to demonstrate clear business value in adhering to EU values¹⁷, and their advantage over existing commercial solutions.
- BC2. **Global Competition:** *Competing in the global market through Product-Service platforms.* The combination of Data and Service Economies represents the major competitive advantage of the EU with respect to global competition. Thus, added value data-driven services that could make ‘Made in EU’ products competitive globally need to be identified. SMEs (99% of the EU industrial fabric) and the role of PPP mediators like the Digital Innovation Hub (DIH) need to be further clarified, and co-competition models (e.g., those put forward by MOBI¹⁸) need further investigation.
- BC3. **Dynamic Ecosystems:** *Implementing data spaces in dynamic business and data ecosystems.* In the industrial domain, shared data ecosystems must guarantee data producers full control over the access and usage of their data. However, ownership is a difficult concept to legally-define (see related legal compliance challenge, LC2) and there are no clear guidelines or consensus on how to implement data sovereignty (see related technical challenge, TC2) in flexible and dynamic business ecosystems (rather than in static hierarchical supply chains). It is also unclear how next generation peer-to-peer networks can guarantee trust and sovereignty without centralised control.
- BC4. **Dynamic Skills:** *Effects of disruptive technology challenges on the job market.* There are differing views on the exact impact that new data-driven technology and automation will have on jobs and employment. Short-term actions include the re-skilling and up-skilling of personnel. However, in the longer term a complete re-definition of workflows, processes, and human-machine interaction patterns (e.g., ‘collaborative intelligence’ between humans and autonomous systems) is required. The current education system is still not geared towards continuously catering for new and unknown professions.
- BC5. **Digital Transformation:** *Challenging organisational impact of the 6Ps digital transformation model.* Data-driven transformations are needed at the level of products (and services), processes (and

¹⁷ An obvious example is European privacy safeguards, enforced by the GDPR, to develop new products (e.g., Personal Data Wallets) that reward users fairly and by design.

¹⁸ Mobility Open Blockchain Initiative: www.dlt.mobi.

organisations), platforms (and spaces, marketplaces), people (and roles), partnerships (and participatory innovation models), and performance (and data-driven KPIs). Methods and tools to support EU Industry with this transformation are required. Although disruptive innovation models benefit start-ups and scaleups, evolutionary innovation models need to be considered as alternatives for the wider industrial ecosystem.

- BC6. **Trust:** *Lack of data sharing trust and motivation.* Data marketplaces rely on an understanding of the commercial value of data produced by industry at all levels. The lack of confidence in the quality of data available for sharing is in itself a challenge. Without quality standards, widespread, automatic data exchanges will not materialise. Attempts at optimising data accuracy should also extend to algorithms (e.g., algorithm bias). In addition, preparation costs for sharing data (e.g., cleaning, quality assurance) need to be considered, together with risks (e.g., potential access to trade secrets, Intellectual Property sharing). In addition, sharing personal data in B2B applications needs to strictly comply with the GDPR. The effective application of anonymisation and obfuscation methods can minimise both risks. However, it can generate increasingly synthetic data, whose classification can be objective and whose actual value can be extremely low (e.g., for critical applications), thus not offering producers an immediate return on investment. Open Data Models can become a new model for industry, if the value of open innovation, early involvement of customers, and strategic alliances (even with competitors, as in the case of service ecosystems) are well understood. To set up trusted data networks, ad-hoc and on-the-fly B2B data exchange mechanisms and contracts, provided under well-defined data sovereignty principles, need to be considered.
- BC7. **Valuation Standards:** *Lack of data valuation standards in marketplaces.* Data marketplaces introduce new opportunities and business models at whose centre lies the valorisation, or monetisation, of data assets. New challenges revolve around the pricing of data, e.g., whether this is determined by the producer, the market demand, or by a broker or third party; whether the value for a specific data asset is universal or depends on the buyer-seller relationship. Guidelines and pricing models need to be established in order to guide business in determining value in participation (refer to last business challenge). New forms of value creation uncovered by new sharing mechanisms need to be explored¹⁹. In addition, data discovery will need to be better addressed since the value of data assets is materialised upon findability.

Legal Compliance Challenges

All the different regulations²⁰ introduced in the last decade in the context of the Digital Single Market make for a complex landscape of policy for data. Notwithstanding the inherent complex role of data, an increased understanding is needed about how data regulation interplays and connects within data platforms. The following are the most pressing challenges that need to be addressed:

- LC1. **Data Protection:** *Tackling inverse privacy and understanding personal data rights.* There is a significant gap between the rights introduced by the GDPR (and its 28 national implementations) and the average citizens' and companies' understanding of their implications, what control they can exercise and how²¹. New business models should not assume that a sufficient portion of private

¹⁹ Token-curated Registries are examples of emerging protocols and networks to address this issue at scale: <https://medium.com/@ilovebagels/token-curated-registries-1-0-61a232f8dac7>.

²⁰ Refer to an overview and discussion under Annex 2, Data Sharing Ecosystems for Government and Public Bodies

²¹ The Cookie Law failure is a recent example of how user control and informed consent have pitfalls

users have the time, expertise and interest to fully comprehend these implications, but also that data rights and consent can change. The practice of inversely private data should be discouraged so as to provide means for individuals to retain control and access to their own private data at all times. More guidance is needed from regulators and data platforms developers. Developments such as sticky policies and dynamic user consent offer promising avenues (also refer to the related technical ‘Privacy Protection’ challenge).

- LC2. **Free-flowing data:** *Ownership and other legal blockers.* Although we speak of the free movement of data as a 5th European freedom, data is far from flowing freely. Legal questions surrounding data ownership, access, portability, retention, etc., remain pressing topics of attention, even more so in an AI context. Existing legislation (e.g., database rights) are outdated, hampering the use of data in AI and the development of new business models. In a data marketplace setting, data ownership is particularly hard to address, as it cannot be legally defined. In the absence of a ‘GDPR for non-personal data’, the principle of Data Sovereignty can be an answer to confidentiality and security requirements, but also poses implementation challenges (see related technical challenge, TC2).
- LC3. **Privacy Preservation:** *Privacy preservation in an open data landscape.* Open data initiatives and public blockchains are driving open innovation in multiple ways. Privacy preservation in this openness is a topic that has to be carefully examined not only in technical terms but also regarding legal compliance in national and European level.
- LC4. **Regulatory Compliance:** *General uncertainty around data policies.* Data-driven SMEs and companies that aim to develop data platforms still face questions on how to incorporate and adjust for the effects of the regulatory landscape within the Digital Single Market; e.g., how to be compliant, when, where and which regulation comes into effect, how to gather knowledge on implementing the regulation, etc.

National and Regional Challenges

Industry and academia are adopting new and disruptive technology much faster than member states and the European Commission can adapt their policies and regulations. In the midst of an emergent data economy facilitated by the convergence of digital technologies these challenges need to be high in the political agenda:

- NC1. **Workforce Skills:** *Public organizations lack digital skills and resources.* Digital technology is developing fast and public organizations have difficulties in keeping up with the pace of development (perhaps more so than business, see also business challenge BC4). At the same time, it is difficult to identify what kind of new skills and education public organizations would need. For instance, new digital skills include skills to plan how data is organized and served in order to create value in society. Organizational and individual skill development are also budget issues, which may not be high on the public agenda. The challenge is to use funding wisely and to avoid waste of resources.
- NC2. **Resistance to change:** *Insufficient support for digital transformation in business by public authorities.* Digitization will transform processes, and data, along with AI, will be used to build up knowledge on society. Transforming the organization leads to changing personnel’s work profiles. Roles will change leading to employment disruptions and the need for re- and up-skilling. New services are an opportunity but resources for the transformation are limited. Efficiency and transparency need data sharing but also investments in order to create new data spaces.

- NC3. **Investment Evaluation:** *Evaluating public organization efficiency and economic impact in data era.* Public organizations serve the society, both citizens and industry alike. The constant requirement of efficiency and impact improvement motivates governments to find out new services based on data. Decisions on development investments are difficult to make, however, and quite often investments are considered risky. Nevertheless, public organizations and their services are an important part of society, and one of the starting points of emerging data ecosystems. From a governmental point of view the challenge is to evaluate investment to data-centric organizations and ensure that economic results have an impact on the whole society.
- NC4. **EU-wide policies:** *Lack of common innovation policies.* Stepping up from regional innovation policies to EU level comparisons is challenging. Data provides a mean to measure impact of innovation policies, but regions find it difficult to compare due to varying requirements. For instance, simple dataset timescale variation may give odd results depending on the region.
- NC5. **Policy Compliance:** *Translating European-wide policies into tangible measurements.* To enable the possibility of real-time, data-driven policy compliance verification, further investments in infrastructure and the certification of data, from devices such as IoT appliances and edge nodes, are required. When data is needed as evidence for compliance with specific regional and European policies, standard or common approaches that are recognised and accepted by the respective policies are required to map data, e.g., from IoT device measurements, into compliance levels.

ABOUT BDVA

The Big Data Value Association (BDVA) is an industry-driven international not-for-profit organisation with over 220 members all over Europe and a well-balanced composition of large, small, and medium-sized industries as well as research and user organisations. BDVA is the private counterpart to the European Commission to implement the Big Data Value PPP program. BDVA and the Big Data Value PPP pursue a common shared vision of positioning Europe as the world leader in the creation of Big Data Value. BDVA is also a private member of the EuroHPC Joint Undertaking and one of the main promoters and driving forces of the AI, Data and Robotics Partnership planned for the MFF 2021-27.

The mission of the BDVA is “to develop the Innovation Ecosystem that will enable the data-driven digital transformation in Europe delivering maximum economic and societal benefit, and, achieving and sustaining Europe’s leadership on Big Data Value creation and Artificial Intelligence”. BDVA enables existing regional multi-partner cooperation, to collaborate at European level through the provision of tools and know-how to support the co-creation, development and experimentation of pan-European data-driven applications and services, and know-how exchange.

BDVA maintains and fulfils a Strategic Research and Innovation Agenda (SRIA) for Big Data Value domain, contributes to the Horizon 2020 work programmes and calls for proposals and it monitors the progress of the BDV PPP (BDVA is in charge of producing the Monitoring Report of the whole programme). BDVA manages over 25 working groups organised in Task Forces and subgroups, and tackling all the technical and non-technical challenges of Big Data Value. BDVA has also developed, together with CLAIRE, ELLIS, EurAI and euRobotics, the third release of the SRIDA (Strategic Research, Innovation and Deployment Agenda) for the European AI, Big Data and Robotics Partnership.

For further information: www.bdva.eu / info@core.bdva.eu / [@BDVA_PPP](https://twitter.com/BDVA_PPP)

Editors: Simon Scerri (Fraunhofer), Tuomo Tuikka (VTT), Irene López de Vallejo (BluSpecs)

Major contributors (in alphabetical order by surname): Martina Barbero (BDVA), Arne Berre (SINTEF), Davide dalle Carbonare (Engineering), Oscar Corcho (UPM), Edward Curry (Insight Centre for Data Analytics), Valerio Frascolla (Intel), Ana García Robles (BDVA), Robert Ginhör (Know-Center), Sergio Gusmeroli (POLIMI), Allan Hanbury (TU Wien), Jim Kenneally (INTEL), Antonio Kung (TRIALOG), Till Christopher Lech (SINTEF), Antonis Litke (NTUA), Irene López de Vallejo (BluSpecs), Brian Quinn (INTEL), Dumitru Roman (SINTEF), Simon Scerri (Fraunhofer), Harald Schöning (Software AG), Tjerk Timan (TNO), Tuomo Tuikka (VTT), Theodora Varvarigou (NTUA), Ray Walshe (DCU), Walter Weigel (HUAWEI).

Other comments from BDVA members and external communities were also received along the process.

Steps in the production of this Position Paper (Version 2):

The milestones listed below were complemented by routine presentations in BDVA Board of Directors (BoD) and Activity Group meetings to gather feedback and comments, as well as bi-weekly calls between interested contributors and more frequently at editorial level.

- Collection of feedback from AIOTI community following shared IOTWeek-19 session
- Collection of feedback from IDSA
- Collection of feedback from the EOSC via their survey: <https://www.eoscsecretariat.eu/towards-european-data-sharing-space>
- Collection of feedback from dedicated EBDVF-19 track panellists and audience
- Collection of feedback from the European Support Centre for Data Sharing following EBDVF-19
- Participation and discussion in the EC Roundtable Accelerating a Sustainable European Data Economy, Amsterdam November 2019
- Prioritising Actions following BDVA Data Sharing Survey (2020)²²
- Various Task Force 1, Sub-group 7 (Data Sharing Spaces) online workshops, latest of which took place on 20.10.2020
- Various BDVA Activity Group Meetings in 2019-2020
- November 2nd: Adoption and Publication of the second version of the paper

Note: This document should be referenced as follows:

López de Vallejo, I., Scerri, S., Tuikka, T. (eds) (2020) Towards a European-Governed Data Sharing Space. Brussels. BDVA

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

²² <https://www.bdva.eu/sites/default/files/Preliminary%20Analysis%20of%20Data%20Sharing%20Practice%20survey%20results.pdf>



www.bdva.eu