

TOWARDS A EUROPEAN DATA SHARING SPACE

Enabling data exchange and unlocking AI potential

BDVA Position Paper

April 2019



BDV

BIG DATA VALUE
ASSOCIATION

EXECUTIVE SUMMARY

The successful development and adoption of a pan-European data sharing space will mark a milestone in the evolution of the new data economy. The envisaged macro-level space can incorporate existing vertical, cross-sectoral, personal and industrial data spaces, offering services and experimentation opportunities to all stakeholders. It also needs to adhere to and promote European values, e.g., democracy, privacy safeguards and equal opportunities. These values can become the trademark of European data economy technologies, products and practices. Rather than be seen as restrictive, legislation enforcing these values should be considered as a unique competitive advantage in the global data marketplace. The exploration of fair, secure and trustworthy legal, regulatory and governance frameworks should take place alongside these developments, together with strategies and investments in the development of the European data workforce. 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 its members, have a strong role to play in steering and guiding these activities at European level. This position paper is therefore meant i) 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 ii) to guide public and private investments in this area in the next Multiannual Financial Framework. This paper addresses all the above mentioned actors and considers joint collaboration as the only way forward. The paper extrapolates current trends to highlight **opportunities** in Europe for industry, private users (citizens), research and academia and the public sector, over the next decade. It also identifies and analyses foreseen **technical and non-technical challenges** (business, organizational, legal compliance and skills-related challenges) as well as specific challenges that need to be tackled in a coordinated manner at both national and regional levels.

Finally the paper includes a **list of recommendations for the successful development, implementation and adoption of a European Data Sharing Space** that i) allows new and existing vertical, cross-sectoral, personal and industrial data spaces to interoperate, ii) offers services and experimentation opportunities to all stakeholders, and iii) promotes European values. In particular BDVA recommends:

- To create the conditions for the development of a **trusted European data sharing framework**.
- To **incorporate data sharing at the core of the data lifecycle** for greater access to data.
- To provide **supportive measures for European businesses** to safely embrace new technologies, practices and policies.
- To assemble a **European-wide digital skills strategy** to equip the workforce for the new data economy.

The content of this paper and outcomes of future dialogues will also feed into the Strategic Research and Innovation Agenda for a European Artificial Intelligent Public Private Partnership (AI-PPP SRIA)¹.

¹ <http://bdva.eu/AIPPP-Vision-paper-PressRelease>

Table of Contents

Executive Summary	2
1 Introduction	4
2 Opportunities in the next 10 years	5
2.1 Opportunities for Business	5
2.2 Opportunities for Citizens	6
2.3 Opportunities for Science	7
2.4 Opportunities for Government and Public Bodies	8
3 Foreseen Challenges	9
3.1 Technical Challenges	9
3.2 Business and Organisational Challenges	10
3.3 Legal Compliance Challenges	12
3.4 National and Regional Challenges	12
4 BDVA Recommendations	13
Annex 1. Term Disambiguation: Data Spaces, Platforms and Marketplaces	16
Annex 2. Prominent Examples from the Current Technical Landscape	17
About BDVA	21

1 INTRODUCTION

A socio-technical shift is gaining momentum, converging both regulatory, research and innovation policies as well as technology and data ecosystems towards implementable strategies, frameworks and technical solutions that can sustain an ethical, open and democratic European data economy. Exploiting these trends will present an opportunity to position Europe at the forefront of data and AI innovation. However, to realise this vision concrete actions need to be urgently taken.

In the above context, this position paper responds to the observed need to present Europe with a new metaphor for the emergent Data Economy, which relies on the availability of data as a basis for further innovation and exponential development of technologies, especially trustworthy AI that reflects European values. Rather than only considering data as the new oil, we think of it as a new currency underpinning a new data economy; catalysing the European economy to develop faster and more effectively.

For AI technology to develop further and meet expectations, additional diverse and large datasets need to be made available. The BDVA position argues that data spaces, platforms and marketplaces are enablers, key to unleash the potential of such data. This position is entirely aligned with the European Commission's plan² to foster the development and use of "made in Europe" AI technology, specifically supporting the call for a joint action with Member States to establish "common European data space to make data sharing across borders seamless" and thus enable "large, secure and robust datasets" for AI technology to be developed. Without wide availability of data we cannot develop AI that can compete with that being produced by countries like the US or China. We must foster the development of AI developed with European ethical values at heart³.

This paper also outlines the role that next-generation data platforms, as data-driven complex systems, perform in driving the European Data Economy and enabling full adoption by both business and private consumers' markets, while implementing EU policies and upholding values regarding democracy, privacy protection and equality. Emerging **Data Ecosystems** that enable large-scale data to be securely connected, valorised and shared, rely on three complementary technologies (details in Annex I):

1. Data storage, lifecycle management platforms and protocols, or simply **Data Spaces**, as networked industrial and/or personal data spaces,
2. Next generation data acquisition and processing platforms, or simply **Data Platforms**, as exemplified by the BDVA reference model and its reference implementations⁴,
3. Data sharing and exchange platforms, or simply **Data Marketplaces**, where data is commercialized using Open Data, Monetized Data and Trusted Data sharing mechanisms⁵.

In contrast to other well-known domain-specific initiatives⁶, the BDVA intends to adopt a wide-angle perspective to explore arising challenges and opportunities for future data ecosystems. This includes

² Member States and Commission to work together to boost artificial intelligence "made in Europe", EC Press Release, Brussels, 7 December 2018, http://europa.eu/rapid/press-release_IP-18-6689_en.htm

³ <https://www.euractiv.com/wp-content/uploads/sites/2/2018/12/AIHLEGDraftAIEthicsGuidelinespdf.pdf>

⁴ Refer to the latest BDVA Strategic Research and Innovation Agenda, http://bdva.eu/sites/default/files/BDVA_SRIA_v4_Ed1.1.pdf

⁵ <https://ec.europa.eu/digital-single-market/en/news/staff-working-document-guidance-sharing-private-sector-data-european-data-economy>

⁶ For example, the Digitising European Industry WG2's focus is on next-generation industrial data platforms, focussing on specific vertical sectors. https://ec.europa.eu/futurium/en/system/files/ged/dei_wg2_final_report.pdf

considering the broadest array of data and data producers, and whose value is raised by components and services targeting all kinds of data consumers and areas of application.

To this end, there remains a need to survey the current landscape, to identify which technologies and solutions have the highest potential to further evolve and disrupt the market over the next decade. In this position paper we limit our observations to the most prominent examples from the emerging technical landscape (included in Annex 2), to identify: i) known and foreseen opportunities, ii) known and foreseen challenges that can limit their impact. Based on these observations, we enlist a set of concrete recommendations for the EC that can be materialized during the Digital Europe and Horizon Europe Programmes, 2021-2027.

2 OPPORTUNITIES IN THE NEXT 10 YEARS

As indicated in the previous section, in recent years a marked interest has been observed by major industrial players, national and European legislative institutions and other key stakeholders in:

- **Alignment and integration of established data sharing technologies and solutions**, avoiding reinventing the wheel and with the intention of supporting scale⁷,
- **Architectures, standards, protocols and governance models aiming to unlock data silos**, over which a) fair and secure data exchange and sharing takes place, b) protection of personal data is paramount, c) distributed and decentralised solutions enabling new types of data value chains can be explored,
- **Business models that can exploit the value of data assets** (including through the implementation of AI) bilaterally or multilaterally amongst participating stakeholders that are not limited to industry but also include local, national and European authorities and institutions, research entities and even private individuals.
- **Fostering and accelerating the uptake of data technologies** and the data economy within sectors whose business models are not yet data-driven.
- **Enabling data analytics across a European data sharing ecosystem** that includes research centres, industry, government and multi-national bodies, by leveraging existing pan-European initiatives and networks (e.g., DIH and i-Spaces).

Below, we extrapolate current trends to outline opportunities arising over the next decade for the four primary societal spheres of industry (Business); private users (Citizens as customers); research and academia (Science) and local, national and European government and public bodies (Government). However, the identified opportunities have a wider socio-economic value, often benefitting multiple spheres, and should therefore not be considered solely within the confines of these categories. Indeed, possibilities to boost the data economy will lead to an improved general economy thus benefiting society as a whole.

2.1 Opportunities for Business

Both SMEs and large industry stand to gain from the following emerging opportunities:

⁷ Overview on established reference frameworks and standardization efforts: <https://i40-tools.github.io/StandardOntologyVisualization/>

- **Open data marketplaces that level the playing field for industrial data sharing.** The emergence of large-scale marketplaces whose participation is open to all kinds of data producers and consumers will complement the, currently dominant, exclusive data sharing agreements and solutions designed for specific parties. Thus, industrial data can be shared inside and outside of a value network, guaranteed that the producers retain data sovereignty and receive fair compensation. European industrial players of all sizes, who provide services that are both tangible or immaterial, are able to tap into data sources that are made available within the rules set by the data producers.
- **Increased availability of vast and heterogeneous data ecosystems for AI.** Advanced AI applications play a fundamental role in the most critical business processes of several industries. Availability of valuable datasets is key for unleashing the potential of AI gross added value, especially in key industries like Business Services, Manufacturing, Wholesale and Retail⁸ and infrastructure providers (e.g., 5G operators). Emerging data marketplace infrastructures enable data to be shared and traded in a trusted, secure and transparent manner that respects ownership. Blockchain technology enables data sharing at scale. Federated analytics on distributed data sources enable the sharing of derived insights without the need to share the original data; thus balancing privacy, autonomy and IP protection⁹. These developments will make data available for AI development in new ways, offering unprecedented opportunities to scale the ecosystem.
- **Innovative data-driven business models enabled by new value ecosystems.** Moving from 'Data for Business Intelligence' to 'Data for AI' also means a focus transition from internal processes to more collaborative and participative cross-domain ecosystems. The most innovative data-driven business models are showing a wide variety of value creation possibilities, from direct data monetization to access-based valorisation of data assets on sharing platforms. The mobilisation of data from business processes and services can offer a myriad of new opportunities, where data-driven often also implies engagement with strategic partners and the development of data ecosystems around these opportunities¹⁰.
- **Opportunities to tap into 'safe' personal data.** The GDPR sets clear requirements for the safe use of de-personalised data outside of its original purpose. Following clear consent and true anonymization (an outstanding challenge) it will increasingly be considered for cross-sectoral applications. Driven by the private customers' use of digital services, cross-sectoral services will continue to emerge. The banking industry, for example, was forced to rethink its market position following the Payment Services Directive and, based on transaction data, new services are being provided across insurance, banking and health service providers.

2.2 Opportunities for Citizens

European citizens will benefit from easy and secure data sharing in various ways:

- **Full control over personal data.** Under the GDPR, data platforms must guarantee legally-compliant data privacy and sovereignty, affording individuals higher control and traceability of their data. In this ideal scenario, individuals will be able to monitor what data they share, which space it is stored

⁸ The three industries contribute for nearly 50% of Value Added, 6,600,000 €M in 2017, according to IDC in DataBench project D2.2 under publication, www.databench.eu.

⁹ An example of this approach is behind the Collaborative Cancer Cloud: www.intel.com/content/www/us/en/cloud-computing/ohsu-precision-medical-analytics-video.html

¹⁰ A prominent example is the drive by the Mobility Open Blockchain Initiative's (MOBI, www.dlt.mobi/) consortium to accelerate the data sharing for autonomous vehicle technology.

in, and who can access or use it; while retaining the right to alter these decisions¹¹. However, in addition to the management of inversely private data, which remains a challenge in itself, the need for individuals to fully comprehend the implications of these rights remains. The BDVA is active in gathering and disseminating cutting edge developments in the area of privacy-preserving technologies for Big Data, contributing to a better digital single market and increased end-user data protection.

- **Well-being and Quality of Life benefits from personal data sharing in key sectors.** GDPR-compliant data sharing platforms enable emerging European technology to perform data analysis for a clear personal (or social) benefit in important sectors such as Health. The safe management of legally-compliant personal health data records¹² allow for broader analysis (see 'data for research' opportunity, next section) of health, wellness and life data for improved understanding, risk identification and prevention of diseases directly benefiting private users¹³.
- **Access to personalised and cross-sectoral B2C services.** Digitisation brings production closer to consumers, offering more choice and personalisation independent of geographic location. Interoperable data ecosystems are required to enable bundling services during digital transactions. The business opportunity introduced by tapping into personal data will therefore also benefit private citizens, e.g., banking and insurance services offering consumers alternative product sales and choices between the most cost-effective options.
- **Increased opportunities of personal data monetisation.** New European legislation incentivises individuals to share their data, introducing C2B business models that allow them to remain in control of their data while directly receiving fair monetary or economic benefits.
- **New professional opportunities.** Further innovation will define new career pathways and generate additional jobs whose education, re-skilling and up-skilling will continue being supported by national and regional authorities that understand their long term value.

2.3 Opportunities for Science

Academia is expected to benefit from the following opportunities:

- **Increasing socio-economic impact of research data across domains and borders.** Converging standards for data and metadata representation, sharing models, licensing and exchange protocols will make it increasingly easier to discover, integrate or otherwise jointly process and analyse data in other scientific domains. This broadens collaboration opportunities between different scientific fields and also promotes the value in making generated data available, at least partially, as open data, for the wider good. In addition, the establishment of appropriate guidelines can promote collaboration between scientific and national bodies¹⁴ to better address societal challenges, effectively reducing data access bottlenecks faced by European researchers.
- **Advancing science and open innovation through data availability.** Access to data for research purposes remains limited since companies need to invest effort preparing data for little or no apparent gain. After clear business incentives for data exchange and digital infrastructure that

¹¹ Examples are <http://mydata.org> and initiatives such as HAT <https://www.hubofallthings.com/> a new, fully scalable and advanced technology that confers intellectual property rights of personal data to individuals through their ownership of a dedicated database, wrapped with containerised microservices.

¹² A prominent Finnish example: <https://www.kanta.fi/en/my-kanta-pages>

¹³ An example from Singapore targeting the early diagnosis of Parkinson's disease: <https://connectedlife.io/>

¹⁴ An example is the data collaboration model established by the Scottish Federated Safe Haven network, supported by the UK government: http://www.nhsresearchscotland.org.uk/uploads/tiny_mce/The%20Scottish%20Federated%20Safe%20Haven%20Network.pdf

removes friction in the process are in place, little additional effort will be required to make the data available (different license agreements) to academia. In return, it will be easier to involve researchers with access to this data in open innovation activities, thus accelerating innovation in companies. Due to the data access conditions, data 'challenges' remain popular in academia¹⁵. However, while platforms such as Kaggle¹⁶ are very successful in this regard, they tend to be limited to specific challenges with little flexibility in the evaluation metrics. An increase in data availability enables academics to design and run more complex challenges, thus improving scientific crowdsourcing to advance science and identify solutions benefitting industry. This can help even out the imbalance between European and non-European industry when it comes to provisioning data to scientists as a core part of their business¹⁷. Aside from industry data, science will also benefit from the GDPR-enabled availability of de-personalised 'personal' data from millions of European citizens, offering new research opportunities in societal challenges such as healthcare (refer to 'Real-time Statistics' point in the section below), transport, etc.

- **Monetisation opportunities brought about by emerging data-driven business models.** Providing controlled access to research data will enable scientists, universities and research institutes to have the opportunity to exchange or monetise their research data by making it available in a controlled way to other institutes and companies. This will strengthen cooperation between research and industry, enable different data to be integrated and analysed and thus introduce new revenue opportunities for academia.

2.4 Opportunities for Government and Public Bodies

These opportunities will benefit all levels of government, national and European public authorities:

- **Data commons for better government services.** The opening up of public domain datasets and systems offers opportunities for the improvement of existing services, or development of new ones. Moreover, it can increase accessibility and simplification of e-Services. The SDGR¹⁸ promises to make company- and citizen data available across Europe in a one-stop-shop manner¹⁹. Harmonising and opening up government data also enables governments to act as data platforms offering digital services to both companies and citizens.
- **AI-enhanced digital services.** AI-boosted digital services can help predict and analyse national and European data in a privacy-preserving and ethical manner. Collaboration, piloting and information sharing between governmental agencies supports data platform exploitation. EU government bodies can form an information sharing platform to support emerging EU-wide Data Service Management, based on working national examples²⁰.
- **Real-time European statistics.** An integrated European data sharing space can provide real-time monitoring across key sectors at both national and EU level. Examples include economy, security, and health: customs statistics can be observed to monitor imports and exports, and also help with cross-border security; whereas health data can make disease outbreaks visible to all relevant healthcare authorities.

¹⁵ An example is the ImageNet Large Scale Visual Recognition Challenge: <http://www.image-net.org/challenges/LSVRC/>

¹⁶ <https://www.kaggle.com/>

¹⁷ Examples from the US and China in this regard include Google, Amazon, and Alibaba.

¹⁸ The Single Digital Gateway Regulation https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2018.295.01.0001.01.ENG, see also <http://europa.eu/youreurope/index.htm>

¹⁹ The once-only principle makes it easy for companies to settle and set up businesses across Europe: <http://toop.eu/>

²⁰ Finland has a network of government agencies to share best practices of introducing AI to the organisations

- **Lean business environment enabled by access to government services.** Public governmental services can be connected with industrial data for leaner business planning. Financial planning can be based on real-time information of the effects of rules and taxation regulations. Better integrated information systems enable automation of taxation, leading to cost saving and predictable national budgets.
- **Evidence-based policy making.** Data for policy, sometimes also referred to as evidence-based policymaking, is the process of using big data in the policymaking process. It allows both policymakers and governmental bodies to use from public sector data repositories, as well as collaborating with private actors to improve and speed up policy cycles and to explore new areas of policy-making in a data-driven way.
- **Data as evidence of Policy compliance.** Policymakers and regulators will increasingly be dependent on data-driven solutions for policy compliance. Traditional compliance mechanisms are being challenged in a data-driven society, due to an increasing velocity and complexity of regulating cyber-physical systems. Advances in the fields of Open Data, Internet of Things and Edge computing create a wide array of new data to be used by public organizations, smart cities, manufacturing and production lines, logistics etc. These data can serve as evidence for validation of whether specific policy-related conditions, implied by European wide regulation and policies, are met. This can open new horizons on how certifications on organizational procedures can be provided. Areas of automating compliance and sticky policies²¹ can already be witnessed in the use of APIs or Blockchain for smart contracting in data markets, for example.

3 FORESEEN CHALLENGES

We include an overview of the most critical challenges threatening value generation introduced by the opportunities outlined. The subsections below offer different perspectives to these challenges but should be considered neither exhaustive nor in isolation. In addition, the fact that some challenges are more prominent in specific sectors is well known and should also be taken into consideration.

3.1 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. The impact of known challenges (e.g., the Vs of Big Data) 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:

- **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

²¹ Pearson, S., & Casassa-Mont, M. (2011). Sticky policies: An approach for managing privacy across multiple parties. *Computer*, 44(9), 60-68.

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.

- **Managing and respecting data ownership.** The realisation of a mixed data sharing space will only materialise if data producers are guaranteed to retain their rights as the original owners. However, data marketplaces rely on a form of ownership that is transferable, making the legal concept of data ownership difficult to define. To safely and fairly support this concept technically, different ownership models or suitable data rights management frameworks need to be explored.
- **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 as well as velocity 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.
- **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.
- **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 the board.
- **Maturity of privacy-preserving technologies for big data.** Although technical solutions for secure and trustworthy data sharing are in place (e.g., Privacy-Enhancing and Privacy-Preserving Technologies) and in continuous development, their uptake is lagging. More flexible ways of allowing uptake of compliance solutions need to be explored.

3.2 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:

- **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 values²², and their advantage over existing commercial solutions.
- **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

²² 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.

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 further investigation, and co-opetition models like those put forward by MOBI²³ should be explored and encouraged.

- **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 technical challenge). In the absence of a 'GDPR for non-personal data', the principle of Data Sovereignty can be an answer to confidentiality and security requirements. Yet, there are no clear guidelines or consensus on how to implement data sovereignty 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.
- **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 current 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.
- **Challenging organisational impact of the 6Ps digital transformation model.** Data-driven transformations are needed at the level of products (and services), processes (and organisations), platforms (and spaces, marketplaces), people (and roles), partnerships (and participatory innovation models), performance (and data-driven KPIs). Methods and tools to support EU Industry with this transformation are required. Although disruptive innovation models benefit startups and scaleups, evolutionary innovation models need to be considered as alternatives for the wider industrial ecosystem.
- **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, IP). 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.
- **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

²³ The three industries contribute for nearly 50% of Value Added, 6,600,000 €M in 2017, according to IDC in DataBench project D2.2 under publication, www.databench.eu.

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 explored²⁴. In addition, data discovery will need to be better addressed since the value of data assets is materialised upon findability.

3.3 Legal Compliance Challenges

All the different regulations²⁵ 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:

- **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 how²⁶. New business models should not assume that a sufficient portion of private 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.
- **Lack of trust in data sharing.** Both in technical and regulatory terms, trusted and secure forms of data sharing between parties (B2B, B2C, B2G) remains challenging. It also provides an opportunity for big data platform developers: to explore technical means (MPC, APIs, Blockchain) to develop legally compliant, smart data platforms.
- **Legal blockers to free-flowing data.** Whereas data has been coined as the 5th freedom within the EU, this analogy seems not to fit entirely, as data is not yet flowing freely across Europe. Legal questions surrounding data ownership, access, portability, retention and so forth remain pressing topics of attention, even more so in an AI context.
- **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.
- **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.

3.4 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

²⁴ 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>.

²⁵ Refer to an overview and discussion under Annex 2, Data Sharing Ecosystems for Government and Public Bodies

²⁶ The Cookie Law failure is a recent example of how user control and informed consent have pitfalls

economy facilitated by the convergence of digital technologies these challenges need to be high in the political agenda:

- **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. 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.
- **Insufficient support for business digital transformation by public authorities.** Digitization will transform processes, and data, along with artificial intelligence, 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.
- **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 important part of society, they will be starting point or even skeleton to emerging data ecosystems. From government point of view the challenge is to evaluate investment to organizations using data so that economic impact is optimal for the whole society.
- **Lack of EU-wide 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, simply timescale of data set may give odd results depending on the region.
- **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.

4 BDVA RECOMMENDATIONS

We identify the following recommended preconditions for the successful development, implementation and adoption of a European Data Sharing Space that i) allows new and existing vertical, cross-sectoral, personal and industrial data spaces to interoperate, ii) offers services and experimentation opportunities

to all stakeholders, and iii) promotes European values. These recommendations²⁷ could feasibly find their effective implementation alongside the Horizon Europe and Digital Europe programmes.

1. **Create the conditions for the development of a trusted European data sharing framework.** In order to enrol a critical mass of stakeholders to engage in pan-European data sharing, measures need to be taken in all Member States to showcase evidence-based benefits for business, government, science and individuals alike. Common concepts, methods and tools are required to enable the development of trusted technologies, products and services that cater for stakeholders across the diversity of the European Union, while respecting individual rights. Industry should explore new business and innovation models that can exploit these structures, and European policy makers need to align their local and national strategies to support the realisation of this framework. Efforts should pay particular attention to: i) data rights management as a pragmatic solution to guarantee data sovereignty; ii) usability, for inclusive and broader access to data for value creation; and iii) data governance, allowing for overarching European-wide policies and procedures to monitor operations in a European data sharing space.
2. **Incorporate data sharing at the core of the data lifecycle to enable greater access to data.** To enable new data value chain opportunities, research efforts need to converge leading technologies that can overcome current and foreseen challenges²⁸ to enable the development of a cross-sectoral European data sharing space. Guiding principles should be the promotion of data life-cycle management strategies that natively incorporate data sharing events and transactions capabilities at their core. This guarantees the digital continuity of both personal and industrial data products and services. Data value chains handling a mix of personal, non-personal, proprietary, closed and open research data need to be supported. Reference architectures, standardisation and harmonisation efforts need to be promoted at European level²⁹, but at the same time be considered as European precursors for worldwide adoption by closely collaborating with relevant international initiatives³⁰. Future practices should also integrate interoperability, security, quality, privacy and ethics by-design. The impact of this action on existing infrastructures should be monitored closely, so as to avoid pitfalls.
3. **Provide supportive measures for European businesses to safely embrace new technologies, practices and policies.** Ahead of becoming data sharing actors, industry and other actors involved in business initiatives need to be supported with data sharing facilities and environments where new business and innovation models can be safely tested. Existing mechanisms such as the network of European Digital Innovation Hubs (DIH) and the BDVA i-Spaces³¹ should be leveraged by industry for safe experimentation and validation under recognised labels. In addition, new mechanisms and instruments like European-wide Regulatory Sandboxes need to be made available as flexible experimental facilities to incentivise and de-risk the exploration and testing of new business and innovation models enabled by disruptive data sharing technology.
4. **Assemble a European-wide digital skills strategy to equip the workforce for the new data economy.** Educational and training policies need to be revisited holistically across the EU to achieve the required development of relevant data skills. A sustainable European data economy relies not only on the

²⁷ Actions are linked to the technical and non-technical (societal, ecosystem & business models, skills and regulation) priorities outline by the BDVA SRIA. See: http://bdva.eu/sites/default/files/BDVA_SRIA_v4_Ed1.1.pdf.

²⁸ The BDVA PPP is delivering multiple solutions to overcome these challenges, which should be analysed and exploited at a larger scale.

²⁹ See BDVA's response to the European Commission's Communication and Consultation "Building a European Data Economy" BDVA Task Force 5 (Policy and Societal) April 2017), p2-3.

³⁰ For example, standardisation efforts to target include the ISO/IEC 23751 "Data Sharing Agreement Framework" <https://www.iso.org/standard/76834.html> and the new ISO/IEC SG7 "Data Usage" study group.

³¹ i-Spaces can serve as an EU-wide benchmarking tool across the DIH network: <http://www.bdva.eu/i-Spaces>

upskilling and reskilling of the existing workforce, but also on a long-term strategic educational policy which covers all education levels. New generations need to be prepared for dynamic careers that can adapt to accommodate emerging roles and professions that sustain the data economy. These efforts should be complemented with strategies for the development or re-definition of new jobs. Envisioning future scenarios where human-technology collaborative intelligence scenarios become a reality, the future workforce needs to be trained in new abilities, e.g., judging bias or assessing liability in both data and algorithms. The implementation of this action will offset the expected drop in jobs brought about by a rapid concurrence of new and disruptive technologies.

ANNEX 1. TERM DISAMBIGUATION: DATA SPACES, PLATFORMS AND MARKETPLACES

Data Spaces is an umbrella term corresponding to any ecosystem of data models, datasets, ontologies, data sharing contracts and specialised management services (i.e., as often provided by data centres, stores, repositories, individually or within 'data lakes'), together with soft competencies around it (i.e., governance, social interactions, business processes). These competencies follow a data engineering approach to optimize data storage and exchange mechanisms; in this way preserving, generating and sharing new knowledge.

In comparison, **Data Platforms** refer to architectures and repositories of interoperable hardware/software components, which follow a software engineering approach to enable the creation, transformation, evolution, curation and exploitation of both static and dynamic data in Data Spaces.

Although distinct, the evolution of the two concepts goes hand in hand and needs to be jointly considered, and both can be considered as the two faces of the same Data Economy 'coin'. Their complementary nature means that commercial solutions often do not distinguish between the two concepts³². Furthermore, due to the particular requirements for the preservation of individual privacy, a distinction between technology and infrastructures that store and/or handle personal and other data has emerged. The evolution of industrial data platforms (considered key enablers of overall industrial digitization) and personal data platforms (services that use personal data, subject to privacy preservation, for value creation) has continued to follow different paths³³.

Moreover, the concept of **Data Marketplaces** has recently gained ground as an enabler of the Data Economy. In analogy with well-known product-service e-commerce portals, these Marketplaces offer a one-stop-shop where data providers (sellers) and data consumers (buyers) can meet, match and trade their respective data assets and requirements, as well as data algorithms and processing approaches. In doing so Data Marketplaces address the disconnected and non-interoperable nature of multiple existing data spaces, as well as components handling the data (often specifically designed for closed platforms targeting specific applications). This has the effect of softening the boundaries between both, in order to maximize data value extraction. Thus, Data Marketplaces operate on top of Data Spaces and utilize Data Sharing and Data Exchange apps provided by Data Platforms to allow the interoperability and trading of data between different subjects, often but not exclusively as B2B. At the basis of Data Marketplaces there is the concept of valorisation and monetization of Data which could be implemented in open, semi-open or closed Data Economy landscapes.

³² For example, the Siemens Mindsphere platform relies on MS Azure data solutions; the Amazon solutions embed both the EC2 applications (as the platform) and the S3 storage (space) services.

³³ As evidenced by the distinct sub-topics covered by H2020 calls:
<http://ec.europa.eu/research/participants/portal/desktop/en/opportunities/h2020/topics/ict-13-2018-2019.html>

ANNEX 2. PROMINENT EXAMPLES FROM THE CURRENT TECHNICAL LANDSCAPE

The realisation of a virtual, heterogeneous but interoperable European data space as described in this paper is largely already technically feasible. However, as outlined in this paper obstacles remain and hinder a critical mass of data providers from committing to its practical realisation. Emerging protocols that underpin increasingly broader, geography- and sector- independent data sharing ecosystems have the potential to address these challenges, leading to an n-fold increase in data availability and ultimately to an exponential increase in AI opportunities³⁴.

In the absence of a full-scale state-of-the-art survey³⁵, we provide a short overview of the most prominent data sharing ecosystems as they have developed in the same four societal spheres under which we categorised the opportunities: Business, Citizens, Science and Governance. However, it must be underlined that this is neither a rigid categorisation (some of the solutions have since been applied to benefit other stakeholders) nor it is meant to be an exhaustive list.

Prominent Data Sharing Ecosystems for Business

Primary initiatives shaping the European Data Economy in the B2B area are being driven by the International Data Spaces (IDS) Association³⁶, Industrial Internet Consortium³⁷ (IIC), Data Market Austria³⁸, Ocean Protocol³⁹ and the IOTA⁴⁰ Foundation. The reference architecture model⁴¹ for an International Data Space (IDS) promoted by IDSA, has its roots in the Industry 4.0 initiative, as suggested by the original 'Industrial' labelling. However, the model has since evolved to apply to broader business data transactions. It is increasingly recognised by key industry as the *de facto* standard for a trusted decentralised data space within which data can be securely shared while respecting sovereignty. A number of private reference implementations are in place. The IIC's Layered Databus⁴² is another emerging standard advocating the need for data-centric information-sharing technology to enable data market players to exchange data within a virtual and global data space. Since late 2018 the IIC joined forces with OpenFog, another relevant consortium⁴³ that provided its own reference architecture. Data Market Austria is an Austrian lighthouse project developing the technical, legal, and business basis for a data market. The technical basis focuses on decentralised provision of both data and services through making standardised metadata visible to a search engine. The legal aspects have been analysed from the Austrian perspective, but this could serve as a basis for similar analyses throughout Europe. Ocean Protocol is a decentralized data exchange protocol that provides governance, incentives and value creation tools enabled by asset tokenization propelled by blockchain technologies⁴⁴. It allows data to be shared and traded in a safe, secure and transparent manner; democratising access to data while retaining privacy rights. Its reference marketplace allows access to

³⁴ BDVA AI Position: <http://www.bdva.eu/sites/default/files/AI-Position-Statement-BDVA-Final-12112018.pdf>

³⁵ Under consideration as a future BDVA publication.

³⁶ <https://www.internationaldataspaces.org/>

³⁷ <https://www.iiconsortium.org/>

³⁸ <https://datamarket.at/>

³⁹ <https://oceanprotocol.com/>

⁴⁰ <https://www.iota.org/>

⁴¹ Latest version 2.0 at the time of writing available at <https://www.internationaldataspaces.org/publications/ids-ram2-0/>

⁴² https://www.iiconsortium.org/IIC_PUB_G1_V1.80_2017-01-31.pdf

⁴³ <https://www.openfogconsortium.org/>

⁴⁴ <https://oceanprotocol.com/>

additional data marketplaces. The IOTA Data Marketplace⁴⁵ is a project of the IOTA foundation. It uses a distributed Ledger architecture called Tangle⁴⁶. Older landmark contributions in this space include the Architecture Reference Model for IoT⁴⁷, whose ambition was to ‘connect vertically closed systems, architectures and application areas so as to create open interoperable systems and integrated environments and platforms’.

Additional initiatives worth mentioning include the iShare⁴⁸ scheme for trusted data exchange in the transport and logistics sector in the Netherlands, and two separate initiatives targeting the Japanese manufacturing industry: the ‘open edge’ computer software platform promoted by the Edgecross⁴⁹ consortium and the Reference Architecture promoted by the Industrial Valuechain Initiative (IVI). Commercial examples include products from Data Republic⁵⁰, Dawex⁵¹, DEX⁵², Veracity⁵³, COGNITE⁵⁴ and UXP⁵⁵. The consistent underlying theme behind these solutions is a promise to maximise value generation for data owners while guaranteeing both safe and secure sharing and the retention of control over their data. Larger initiatives that offer a competitive challenge to European initiatives include the DataTransferProject⁵⁶, backed by strong US players and promising users online data portability, and the MadeInChina2025⁵⁷ strategy for the digitisation of production centres across the manufacturing industry.

Prominent Data Sharing Ecosystems for Citizens

Citizen-data is often collected by companies, and may, in anonymised form, be part of the B2B activities described above. However, there are initiatives with the aim of getting citizens to take control of the valorisation of their data. Self-sovereign identity⁵⁸ management tools and project are emerging, in which the idea is to reverse the business logic surrounding ownership, control and access rights to data, and to put the end-user, and citizens centre-stage⁵⁹. Besides a new area for digital identity management, there is evidence of opening up data for citizens and there exist platforms for data sharing among citizens, between citizens and government and/or citizen-led data sharing with companies.

A prime example of technical developments aiming to directly benefit private users is MyData⁶⁰, a human-centred personal data management approach that combines industrial need for data with digital human rights. Individuals can choose to use their data to develop communal knowledge for making informed decisions, and to interact more efficiently both with each other as well as with organisations. The International Human Accounts Network aims to build the foundation for a fair and functioning data

⁴⁵ <https://data.iota.org/>

⁴⁶ <https://blog.iota.org/iota-data-marketplace-cb6be463ac7f>

⁴⁷ <https://link.springer.com/book/10.1007/978-3-642-40403-0> produced in the context of the IoT-A FP7 lighthouse project <https://cordis.europa.eu/project/rcn/95713/factsheet/>

⁴⁸ <https://www.ishareworks.org/en>

⁴⁹ <https://www.edgecross.org/en/>

⁵⁰ <https://www.datarepublic.com/>

⁵¹ <https://www.dawex.com/en/>

⁵² <https://www.dex.sg/>

⁵³ <https://www.veracity.com/>

⁵⁴ <https://www.cognite.com/>

⁵⁵ <https://cyber.ee/products/secure-data-exchange/>

⁵⁶ <https://datatransferproject.dev/>

⁵⁷ <http://english.gov.cn/2016special/madeinchina2025/>

⁵⁸ <https://medium.com/evernym/7-myths-of-self-sovereign-identity-67aea7416b1>

⁵⁹ https://medium.com/@timberners_lee/one-small-step-for-the-web-87f92217d085

⁶⁰ <https://mydata.org/>

economy⁶¹ by creating a data exchange method and establishing European-level rules and guidelines for ethical data use. Aurora⁶² (Finland) is an artificial intelligence and autonomous applications network that helps create the conditions for a people-oriented, proactive society. The valorisation of one's personal data is also behind commercial products like DataCoup⁶³, which promise citizens the possibility to take back control of their personal data, monetize it and profit from its sharing on a voluntary basis (C2B). On the technical and policy side, the e-sides⁶⁴ project is an EU-funded project that explores societal and ethical implications of big data technologies and provides a broad basis and wider context to validate and monitor privacy-preserving technologies. The DECODE project⁶⁵ aims to ensure people's ownership of their personal data, providing tools that give them the choice to either keep their data private or share it for the public good. The Hub of All Things⁶⁶ (HAT) is the first ecosystem to fully resolve the technical, legal and commercial barriers in ownership and transfer of personal data between individuals and organisations. HAT is already referenced globally as a best in class solution in the emerging personal data exchange market.

Prominent Data Sharing Ecosystems for Science

Although business application possibilities resulting from these initiatives abound, the main focus remains on enabling more opportunities for research. The European Open Science Cloud⁶⁷ (EOSC), in implementation phase until 2020, is promoted as a cloud for research data in Europe. Although the main objective is to facilitate the exchange of research data between different scientific communities, the initiative also explores new business models seeking to make available scientific data to industry. The EOSC complements H2020 e-infrastructure⁶⁸ activities seeking to establish an open European space for on-line research. Various projects funded to meet this objective focus on specific scientific areas, which also offer data repositories and data sharing capabilities. For example, Elixir⁶⁹ focusses on life science data, including a Data Platform to identify data sources, and a Compute Platform to store, share and analyse datasets. RISIS⁷⁰ has a datasets portal for research and innovation policy studies. The Human Brain Projekt Neuroinformatics Platform⁷¹ allows neuroinformatics data to be uploaded, managed, cited, and shared (although data upload must be applied for and approved).

There are also initiatives that are broader in scope, envisaging a data sharing platform that is not restricted to any one domain or community. The FAIR principles⁷² are recognised as guidelines for facilitating data sharing across diverse stakeholders. The National Research Data Infrastructure is an example of a national initiative by the German government which seeks to establish research data spaces for various domains under one umbrella⁷³. Zenodo⁷⁴ is a platform, hosted by CERN, for archiving and sharing research data (in all formats) from all fields of science and the humanities. All publicly available uploads are assigned a

⁶¹ <https://www.sitra.fi/en/topics/fair-data-economy/>

⁶² <https://vm.fi/ihmiskeskeinen-yhteiskunta>

⁶³ <http://datacoup.com/>

⁶⁴ <https://e-sides.eu/>

⁶⁵ <https://www.decodeproject.eu/>

⁶⁶ <https://www.hubofallthings.com/main/what-is-the-hat/>

⁶⁷ <https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud>

⁶⁸ <https://ec.europa.eu/programmes/horizon2020/en/h2020-section/e-infrastructures>

⁶⁹ <https://www.elixir-europe.org/>

⁷⁰ <https://www.risis2.eu/>

⁷¹ <https://www.humanbrainproject.eu/en/explore-the-brain/share-data/>

⁷² Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., ... & Bouwman, J. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data*, 3.

⁷³ https://www.dfg.de/en/research_funding/programmes/nfdi/index.html

⁷⁴ <https://zenodo.org/>

Digital Object Identifier to make them easily referenceable. Multiple license models for the data are supported. Dataverse⁷⁵ is an open source research data repository software that currently has 39 installations worldwide. The European Data Infrastructure's⁷⁶ (EUDAT) vision is to share and preserve data across borders and research communities through a Collaborative Data Infrastructure. Directories of research-data repositories include re3data.org and FAIRsharing⁷⁷.

A number of other examples seek to facilitate data sharing between established scientific communities in specific fields. The Global Earth Observation System of Systems (GEOSS) and its European contribution, NextGEOSS, are leading examples of data hubs for the EO community, connecting observing systems in order to enable operations on one integrated and coherent data set. Examples from other communities (e.g., DataMed⁷⁸ for biomedical data), abound and are not listed here.

Prominent Data Sharing Ecosystems for Government and Public Bodies

Ever since the emergence of e-Government, data sharing platforms for (inter)governmental data exchange have proliferated. However, many initiatives still face challenges of technical, organisational and/or political nature. Hackdeoverheid⁷⁹ (Netherlands) is one of the first notable initiatives that drove governments to consider providing a data platform for third parties (notably startups) to build services upon. Ever since, (open) data sharing platforms have emerged in different forms and at different speeds in the Member States, some developing their own technical solutions⁸⁰. Besides national open data portals, cross-border EU data sharing has been developed in the form of portals (e.g., EU Open Data Portal⁸¹). An example of a cross-border integrated data sharing platform with accompanying services can be seen in the Estonia-Finland collaboration⁸². Besides inter/intra-governmental data platforms and portals, there are also developments towards accessing private data for the public interest⁸³. The EU has recently pushed regulation for the re-use of public sector data⁸⁴. Other initiatives worth mentioning are the Semantic Interoperability Framework⁸⁵, Connecting Europe Facility's building blocks⁸⁶, the Once-Only Principle⁸⁷ and the Single Digital Gateway Regulation⁸⁸, which together shape an ecosystem that allows companies and citizens to easily access data and interact with digital services across member states.

Data sharing for- and by government plays an increasingly important role in city management. The FIWARE smart cities platform⁸⁹ is a prominent example, amongst many others that aim to become an (inter)city 'operating system'⁹⁰. The MSEC EU-Japan collaboration⁹¹ seeks to develop a marketplace of IoT devices and

⁷⁵ <https://dataverse.org/>

⁷⁶ <https://www.eudat.eu/>

⁷⁷ <https://fairsharing.org>

⁷⁸ <https://datamed.org>

⁷⁹ <https://www.hackdeoverheid.nl/>

⁸⁰ For example, <https://e-estonia.com/solutions/interoperability-services/x-road/>

⁸¹ <https://data.europa.eu/euodp/data/>

⁸² https://www.huffingtonpost.com/adi-gaskell/estonia-and-finland-enter_b_14281296.html

⁸³ <https://ec.europa.eu/digital-single-market/en/news/workshop-access-public-bodies-privately-held-data-public-interest>

⁸⁴ <https://ec.europa.eu/digital-single-market/en/european-legislation-reuse-public-sector-information>

⁸⁵ https://ec.europa.eu/isa2/eif_en

⁸⁶ <https://ec.europa.eu/cefdigital/wiki/display/CEFDIGITAL/Building+Blocks>

⁸⁷ <http://www.toop.eu/>

⁸⁸ https://ec.europa.eu/growth/single-market/single-digital-gateway_en

⁸⁹ <https://www.fiware.org/community/smart-cities/>

⁹⁰ see <http://vital-iot.eu/sites/default/files/VITAL-OS.pdf> and <http://themobilecity.nl/> and

<https://medium.com/intelligent-cities/city-os-the-city-operating-system-b8888e579b64>

⁹¹ <https://www.msecproject.eu/>

services for urban environments. The PRIDE project⁹² is an example of a data ecosystem being developed to govern both city and regional sector-specific data (energy) and provide a regional public service integrating lifecycle management for interoperability, security and privacy.

Other efforts look at data to improve or develop new policies. Examples include DataForPolicy⁹³, Readie⁹⁴ and Policy-labs⁹⁵, which involve multi-stakeholder processes to explore evidence-based policymaking.

ABOUT BDVA

The Big Data Value Association (BDVA) is an industry-driven international not-for-profit organisation with 200 members all over Europe and a well-balanced composition of large, small, and medium-sized industries as well as research and user organizations. BDVA is the private counterpart to the EU 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.

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.

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 Lopez de Vallejo (DEX I Ocean Protocol Foundation)

Major contributors: Sergio Gusmeroli (POLIMI), Tjerk Timan (TNO), Antonio Kung (TRIALOG), Allan Hanbury (TU Wien), Robert Ginhör (Know-Center), Antonis Litke (NTUA), Theodora Varvarigou (NTUA), Brian Quinn (INTEL), Jim Kenneally (INTEL), Arne Berre (SINTEF), Edward Curry (Insight Centre for Data Analytics), Ana Garcia Robles (BDVA), Harald Schöning (Software AG), Walter Weigel (HUAWEI), Simon Scerri (Fraunhofer), Tuomo Tuikka (VTT), Irene Lopez de Vallejo (DEX I Ocean Protocol Foundation).

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

Steps in the production of this Position Paper (Version 1)

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.

- June 2018: BDVA BoD launches the activity to create a position paper further elaborating in each of the pillars of the BDVA FP9 position paper

⁹² Platform for Regional Innovation to enable Data exchange on Energy <https://smile-smartgrids.fr/en/projects/en-pride.html>

⁹³ <http://dataforpolicy.org/>

⁹⁴ <https://readie.eu/>

⁹⁵ See Data-Driven Policy Making: The Policy Lab Approach: https://link.springer.com/chapter/10.1007/978-3-319-64322-9_9. See also <https://blogs.ec.europa.eu/eupolicylab/about-us/>

- 12th September 2018: Workshop at Activity Group meeting, AG28, in Brussels (mainly BDVA members)
- 18th September 2018: Data4AI workshop in Brussels⁹⁶
- 10th October 2018: Available materials shared for comments and input
- 14th November 2018: EBDVF 2018 Workshop⁹⁷
- 23rd January 2019: Workshop at Activity Group meeting, AG30, in Brussels
- 25th February 2019: Presented to BoD for discussion, comments and feedback
- 26th February 2019: Workshop at Activity Group meeting, AG31, in Brussels (identifying recommendations)
- 25th March 2019: Version for comments and final feedback sent to the BoD
- 12th April 2019: BoD approval Adoption and Publication of the first version of the paper

Upcoming milestones and editions

- April - June 2019: Additional input from relevant external communities.
- May - October 2019: Additional input from large data owners (industry)
- October 2019: Additional input from selected H2020 ICT13a projects
- Autumn 2019: Additional input from Member States
- October 2019: EBDVF 2019 Workshop
- November-December 2019: *Publication of Version 2*

Note: This document should be referenced as follows:

Lopez de Vallejo, I., Scerri, S., Tuikka, T. (eds) (2019) Towards a European Data Sharing Space. Brussels. BDVA

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

⁹⁶ Involvement of BDV PPP lighthouse projects, euRobotics, DIHs for Robotics, AIOTI, IDSA, EFFRA, ECSO, AI on demand platform.

⁹⁷ <https://www.european-big-data-value-forum.eu/european-data-space-connecting-data-ecosystems/> with participation of the European Commission, IDSA, Boost4.0 project, Safe-SEED project, Austrian Data Market, KU-Leuven and representation from SMEs and large companies.



www.bdva.eu