BDVA position paper

BDVA's response to the European Commission's Data Strategy "A European strategy for data" May 2020

BDVA welcomes the EC Data Strategy as the natural evolution of the data innovation ecosystem within Europe. The Data Strategy leverages the work of the Big Data Value PPP to support Large Industry working together with SMEs and research organisations in a critical mix needed to deliver both economic and societal impact to Europe in areas including mobility, manufacturing, bio-economy, health, energy, environment, smart cities and infrastructures, food, finance, media and farming.

The Big Data Value contractual Public-Private Partnership (BDV cPPP) is playing a key role in enabling the digital transformation of our society and economy, and in implementing the Digital Single Market Strategy where it supports areas that include data technologies and infrastructures, data platforms, data-driven business (models) and innovation, data standardisation, and skills. The Big Data Value Association (BDVA), the private counterpart of the cPPP, sees clear evidence that the BDV cPPP has successfully mobilised key stakeholders and private investment in the European Data Economy. However, we are just at the beginning of the digital transformation of our society, and the potential for future data-driven innovations is increasing. The BDVA community facilitates the essential involvement of a broad range of stakeholder's perspectives and creates the starting points for necessary further cooperation on responsible development of disruptive technologies for the benefit of both business and society.

Europe has a unique ecosystem of diverse languages and cultures, which should be leveraged to showcase leadership in interoperability, inclusivity, diversity in providing technology solutions in complex environments. Standards and standardisation can be employed as a mechanism to leverage international best practice to build trust in Data Sharing products, services, tools and processes.

BDVA **supports further research and experimentation** on the utilisation of data, AI technologies and data-driven innovation for the good of business and society via a smart mix of technical, legal, ethical and business methods. Research efforts should focus on the concepts, the challenges in implementation and the practical aspects of notions as public interest data, the ethics of data sharing, sustainability-by-design and responsible data engineering also trying to investigate the business opportunity of data sharing beyond their monetisation. The nine research and innovation projects funded under ICT13a[1] (Industrial and Personal Data Platforms) are already providing valuable insights across a diverse set of domains on how to address many of the issues mentioned in the strategy that are holding the EU back from realising its data potential: availability of data, data interoperability and quality, data governance, data infrastructures and technologies, empowering individuals, data literacy and cybersecurity.

---

[1] H2020 ICT-13-2019

The digital strategy splits data from AI and from the regulation of online platforms, while recognising that data feeds AI as 'raw' material and lies at the heart of online platforms. To avoid the risk that crucial technological, economic and societal implications of data are overlooked, data should not be regulated as a standalone product in the abstract. As recognised by the OECD,[2] data is not a monolithic asset. Its value depends on the context of its use "with different implications for individuals, businesses and policy makers"; which should be duly taken into account when regulating them.

BDVA acknowledges the recognition, in the Data Strategy, of the **critical role of data governance**. By focusing on the institutional arrangements between the interested stakeholders, data governance mechanisms can open avenues for regulating data, while paying due consideration for the context in which they are processed. Based on case-specific studies, the factual and regulatory factors that influence data governance need further research, to draw lessons for further regulation.

Data governance is a process that binds together the technical, organisational, legal, quality and business aspects of data. Any organisation that uses data in their core business process needs to have some basic data governance process in place; for data-driven organisations data governance is a *condition sine qua non*. The reason why data governance needs to be addressed in an integrated way is that all aspects influence each other. For instance, low-quality data can be the result of an organisation that does not have incentives in place to guarantee good data; it may lead to legal implications (liability), can only partially be solved by technology and will surely affect business in a negative way. While this is true for individual organisations, it is even more true for organisations that collaborate using data. Effective governance stimulates trust and efficiency. Sharing data without having some kind of data governance in place is almost impossible.

Data Quality Assurance: Generating 'sharing ready' high-quality data that can be valued in a more standard and objective manner remains a challenge. Apart from the need to further invest in methods that reduce data preparation costs (cleaning, quality assurance, respecting IP and avoiding revealing trade secrets, etc), quality standards also need to mature in order to guide data providers on ensuring expected levels of accuracy, completeness and consistency are met. Without guaranteed levels of data quality, widespread, automatic data exchanges will not materialise. Attempts to optimise data accuracy also need to extend beyond data to encompass algorithms (e.g., overcoming algorithm bias).

The Data Strategy should put significant emphasis on the **economics of the Data Value Chain** (different actors bringing different value along the value chain) and on **business models**. When scaling (cross) sectoral data sharing, viable business roles must emerge. To store, move, transform and extract value from data has a cost to public and private organisations and this should be clearly reflected in the European Vision and Strategy. There must be a business model for the different roles that are necessary in the ecosystem.

---

[2] OECD, Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-use across Societies, 2020.

Strategy. A cross-sectoral governance framework for data access and use

The European Commission's data strategy considers reinforcing data governance mechanisms to enable data-driven innovation while also endowing European businesses and citizens with 'their' data. The strategy refers to data pools, data trusts, data cooperatives, to what could be added to data commons, data utilities and subcategories of all such mechanisms. The European Commission also considers the regulation of usage rights for co-generated data in the IoT.

BDVA recognises these challenges, specifically when weighing between horizontal and context- or sector-specific regulation of data as assets. **An experimental research approach** is therefore needed to identify the factors for success or failure, e.g. the technology, the nature of data and of stakeholders, the objectives assigned to the governance mechanism, and the legal framework.. It is only against this background that room for further legislative initiatives (e.g. in the Data Act) can be identified. Data forms the "super-ecosystem" that links Cloud, AI, IoT and 5G to deliver the next generation of disruptive solutions for B2B, B2G and M2M data sharing services. Data usage on such a scale requires careful design of standardisation, certification, regulation, legislation and compliance solutions. In regard to data in the B2B area, access and usage of data should be clearly defined. Partners must have the choice and control over the way in which their data is accessed and used in line with predefined, transparent rules. Partners should be able to take their essential data with them at the end of an engagement.

The BDVA community and activities (i.e. on data governance, data standardisation and on data sharing) and the research and innovation projects of the last phase of the BDV cPPP will contribute to this research topic. Interdisciplinary research will scrutinise the different mechanisms being tested throughout the EU, in order to identify (context-specific) factors for success, results that will provide building blocks for the upcoming research investments and the planned High Impact Project. As a reference, the federation of Data-Driven Innovation Hubs and the new generation of European Data incubators (funded under DT-ICT-05[3]) have as their overarching objective the mobilisation, use and sharing of data between sectors and borders, providing different governance frameworks to align needs coming from different sides (offer / demand) and moving towards a Common European Data Space.

The multi-stakeholder approach should more directly extend to research organisations. Although research data needs to be dealt with in a special manner, it should not be considered as a completely new data sphere requiring its own separate ecosystem. Existing activities in the EOSC should be better integrated in the overall data strategy as a subset of the envisioned common European data space, rather than as a parallel 'data silo'. We are convinced that the solution to the post-pandemic situation will be to apply various approaches of Open Science, building upon the "as open as possible, as closed as necessary" principle. We are confident that this approach will solve the current simplistic dichotomy (i.e. open vs close) and ultimately promote an efficient European Research Area. In line with this view, the specific opportunities, challenges and needs of R2B, B2R (Research) should also be specifically and more visibly included.

---

[3] H2020 DT-ICT-05-2020

Strategy. Enablers

The BDVA community recognises the importance of a dynamic and vibrant "innovation ecosystem" linking the research activities in Universities and Research Centres with the development, deployment and standardisation of technologies by Enterprise and SDOs[4]. BDVA welcomes an integrated approach via the High Impact Project on European Data Spaces and reiterates the importance of linkages with DT-ICT-05 and research data-rich projects and platforms that go beyond an exchange of good practice, to form a Europe-wide federation, to foster real international cross-borders data innovation.

Ongoing and future BDV PPP projects could contribute in a very valuable way to the new High Impact Project and the development of the new European data spaces. The portfolio of BDV PPP projects includes a set of tools and methods in data management and processing[5], data privacy and protection[6], data sharing[7], integration of data analytics with other technologies[8], large scale deployments[9], etc, that could be used as the basis for the development of these new initiatives, learning from past experiences and avoiding starting from scratch. More specifically, projects funded under DT-ICT-05-2020 are expected to provide data sharing tools and platforms, data governance frameworks, and availability, quality and interoperability of data, which are supposed to be the pillars of the new European data spaces, so an alignment between all those initiatives would be desirable.

In addition, DIHs included in these projects should be able to strengthen sustainable and open collaborations with different organisations, including large enterprises, creating an ecosystem to share infrastructures, research objectives, talents and networks. New business models, establishing added value for each organisation, should be developed.

Strategy. Competences

BDVA supports the **Data Literacy** initiative as part of the broader Digital Skills and Education agenda. The skills required by industry partners to deliver on the promise of the Data Economy need to be identified, classified and quantified.

Data democratisation requires appropriate skills to be present in all involved parties, for example data providers, analysts and senior decision makers. Tailored training is required for each party regardless of technical background or experience.

Upskilling and retraining through innovative training (MooCs, Short Courses, Webinars, undergraduate and postgraduate programmes by Higher Education Institute and Industrial Training) needs to be considered along with the wider education and awareness-raising of the public through outreach, dissemination of information on Data products and services, myth-busting and protecting the public from fake-news. This may require a separate coordinated action by the European Commission to facilitate a Digital Skills Platform and ensure Data Democratization.

---

[4] Standards Development Organisations
[5] H2020 ICT-14-2016-2017, ICT-16-2017
[6] H2020 ICT-18-2016
[7] H2020 ICT-13-2018-2019
[8] H2020 ICT-11-2018
[9] H2020 ICT-15-2016-2017

There is a narrow definition of **SMEs** in the context of the current Data Strategy, as SMEs can be solution providers and/or data solution consumers, and can, therefore, play many different roles within the data value chain. All relevant viewpoints should be included. In this context, the support from Data-Driven Innovation Hubs and BDVA i-Spaces to SMEs and start-ups iincludeas a key pillar the definition and implementation of training programs (either based on their own experience or in combination with universities of their ecosystems) aimed at closing the gap of companies workforce on data literacy and skills.

Roll-out: Data Spaces

BDVA recognises and acknowledges that Data Spaces essentially exist in a heterogeneous landscape of different technologies, products and platforms. **Innovation in tools and techniques** will be required to facilitate the interoperability of data sharing among these different and diverse platforms. Standardisation has a key role to play here to ensure the broadest engagement of stakeholders in this process[10].

In addition to the vertical Data Spaces identified in the Data Strategy (manufacturing, Green Deal, mobility, health, financial, energy, agriculture, public administration and skills) BDVA suggests the inclusion of a European Data Space on Media. One of the main challenges of the European media industry is to remain a beacon of trust in a world where unverified information is increasingly communicated through other channels, such as social networks. The creation of a shared data space that will facilitate European fact-checking networks in news verification is a crucial building brick which is currently missing. Moreover, smaller European media companies and platforms lack the capacity to gain wider consumer insights in comparison to the larger platforms of the competition coming from the USA. The creation of a European media data space in which European media companies and platforms share consumer data and insights will allow the creation of new advertising strategies and models, keeping the revenue inside the European economy.

BDVA recognises that there will inevitably be **common issues across the many Data Spaces** and to **prevent a silo approach to innovation a horizontal coordination action or space** is necessary for concertation, integration and collaboration of Data Space innovations.

Finally, the BDVA community envisages and supports a natural evolution from European to global Data Spaces for an international approach to data innovation and advancement of the European Data Agenda through a strategic approach to Data Governance and Data Sharing to best help achieve the objectives of the Digital Single Market. International cooperation is needed to deliver on interoperability requirements for the Digital Single Market (Cross Border Trade and Mobility). Standardisation is what service providers need to deliver quality, trust and confidence to their end-users.

---

[10] E.g. In Smart City scenarios, although cities are already making use of standard (or commonly agreed and shared) data models, interoperability has to be pushed forward in terms of re-usability of the data, how one data-set might be re-used from one sector to another in the same city and/or from city-to-city.

About BDVA

The Big Data Value Association (BDVA) is an industry-driven international not–for-profit organisation with over 200 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 JU 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 developed, together with euRobotics, the consultation version of the SRIDA (Strategic Research, Innovation and Deployment Agenda) for the AI, Big Data and Robotics Partnership.

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