D2.1: Report on high level consultation
A (Big) Data Ecosystem in the making

Workpackage: WP2 – IMPACT. FRAMING the European Data Economy to maximize Impact

Editor(s): Bas Kotterink, Tjerk Timan, Somayeh Djafari

Responsible Partner: TNO

Contributors: Sonja Zillner, Siemens; Nuria de Llama, Atos; Richard Munne Caldes, Atos

Status-Version: Final-V.10

Due to: 15/06/2017

Submission Date: 15/06/2017

EC Distribution: Public

Abstract: This deliverable covers the first results of activities planned in task 2.1 and 2.2. It provides a first overview of the structure and dynamics of the Big Data Ecosystem in Europe. Being the first version of a living document that will be updated throughout the project and will lead to one final report. The report covers the following topics:
D2.1: Report on high level consultation

(i) approach to capturing and reporting on ongoing developments in the EU Big Data Ecosystem, (ii) an overview of Big Data Ecosystems and (iii) two specials, one on business and one on the free flow of data. The deliverable is rounded up with a set of key insights and first recommendations for BVD-PPP stakeholders.
### History

<table>
<thead>
<tr>
<th>Version</th>
<th>Date</th>
<th>Modifications Introduced</th>
<th>Modification Reason</th>
<th>Modified by</th>
</tr>
</thead>
<tbody>
<tr>
<td>V0.8</td>
<td>15/06/2017</td>
<td>First Draft</td>
<td></td>
<td>Bas Kotterink</td>
</tr>
<tr>
<td>V09</td>
<td>06/07/2017</td>
<td>Final draft for review</td>
<td></td>
<td>Bas Kotterink</td>
</tr>
<tr>
<td>V0.91</td>
<td>13/07/2017</td>
<td>Reviewed document</td>
<td></td>
<td>Sonja Zillner</td>
</tr>
<tr>
<td>V0.91</td>
<td>13/07/2017</td>
<td>Reviewed document</td>
<td></td>
<td>Ricard Munne Caldes</td>
</tr>
<tr>
<td>V0.8</td>
<td>15/06/2017</td>
<td>First Draft</td>
<td></td>
<td>Bas Kotterink</td>
</tr>
<tr>
<td>V1.0</td>
<td>26/07/2017</td>
<td>Final version for Quality Review</td>
<td></td>
<td>TNO, ATOS</td>
</tr>
<tr>
<td>FINAL</td>
<td>26/07/2017</td>
<td>Final version ready for submission</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
# Contents

## 1 INTRODUCTION

1.1 SCOPE OF THE DELIVERABLE

1.2 DOCUMENT STRUCTURE

## 2 METHODOLOGY

2.1 OVERVIEW

2.2 SOURCES, DATA GATHERING AND CONTENT DEVELOPMENT

## 3 AN EMERGING BIG DATA ECOSYSTEM IN EUROPE

3.1 INTRODUCTION & METHODOLOGY

3.2 ON ECOSYSTEMS, CLUSTERS AND HUBS

3.3 LAW AND POLICY

3.4 MARKET

3.5 ARCHITECTURE AND INFRASTRUCTURE

3.6 NORMS

3.7 SECTORIAL PERSPECTIVE

## 4 BUSINESS SPECIAL – DATA APPLICATION SCENARIOS

4.1 INTRODUCTION

4.2 THE (GENERAL) BUSINESS CASE FOR BIG DATA

4.3 BIG DATA VALUE CHAINS AND ECOSYSTEMS

4.3.1 Some key trends and challenges within DDI business models

4.4 TWO SHORT SCENARIOS: AGRICULTURE AND TRANSPORT

4.4.1 Making crops smart - potential DDI in Agriculture

4.4.2 Self-driving media centres: the changing landscape of transportation

4.4.3 Summary

4.5 HOW TO RESEARCH BUSINESS MODELS IN THE BIG DATA ECOSYSTEM

## 5 POLICY SPECIAL – FREE FLOW OF DATA

5.1 INTRODUCTION

5.2 ENABLING FREE FLOW OF DATA

5.3 FREE FLOW OF DATA AND COMPETITIVENESS

5.4 CHALLENGES IMPLEMENTING THE FREE FLOW OF DATA INITIATIVE

5.5 IMPACT OF FREE FLOW OF DATA INITIATIVE

## 6 CONCLUSIONS

## 7 REFERENCES
D2.1: Report on high level consultation

List of Figures

FIGURE 1: AN EXAMPLE PORTAL ................................................................. 15
FIGURE 2: FOUR REGULATORY FORCES ................................................... 18
FIGURE 3: MAPPING OF ACTORS IN THE GLOBAL BIG DATA ECOSYSTEM .... 21
FIGURE 4: OVERVIEW OF THE EUROPEAN DATA MARKET ....................... 23
FIGURE 5: DRAFT BIG DATA REFERENCE ARCHITECTURE ....................... 24
FIGURE 6: BIG DATA VALUE CHAIN .......................................................... 30
FIGURE 7: STEPS IN THE DATA VALUE CHAIN .......................................... 34
FIGURE 8: MODEL FOR THE BIG DATA ECOSYSTEM ................................. 35
FIGURE 9: TYPES OF BUSINESS MODELS ................................................ 37
List of Tables

Table 1: Definitions, Acronyms and Abbreviations ........................................... 7
Table 2: Overview of Business Model .................................................................... 33
### Definitions, Acronyms and Abbreviations

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>BDES</td>
<td>Big Data, Ethics, and Society</td>
</tr>
<tr>
<td>BDVA</td>
<td>Big Data Value Association</td>
</tr>
<tr>
<td>BDVe</td>
<td>Big Data Value ecosystem</td>
</tr>
<tr>
<td>DDBM</td>
<td>data-driven business model</td>
</tr>
<tr>
<td>DDI</td>
<td>Data Driven Innovation</td>
</tr>
<tr>
<td>DSM</td>
<td>Digital Single Market</td>
</tr>
<tr>
<td>EC</td>
<td>European Commission</td>
</tr>
<tr>
<td>EDVF</td>
<td>European Big Data value Forum</td>
</tr>
<tr>
<td>EU</td>
<td>European Union</td>
</tr>
<tr>
<td>FFOD</td>
<td>Free Flow of Data</td>
</tr>
<tr>
<td>GDP</td>
<td>Gross domestic product</td>
</tr>
<tr>
<td>HPC</td>
<td>High Performance Computing</td>
</tr>
<tr>
<td>IDS</td>
<td>Institute of Development Studies</td>
</tr>
<tr>
<td>IoT</td>
<td>Internet of Things</td>
</tr>
<tr>
<td>OECD</td>
<td>Organisation for Economic Cooperation and Development</td>
</tr>
<tr>
<td>PPP</td>
<td>Public-Private Partnership</td>
</tr>
<tr>
<td>ROI</td>
<td>Return On Investment</td>
</tr>
<tr>
<td>SME</td>
<td>Small Medium Enterprise</td>
</tr>
<tr>
<td>SRIA</td>
<td>Strategic Research and Investment Agenda</td>
</tr>
</tbody>
</table>

**Table 1: Definitions, Acronyms and Abbreviations**
Executive Summary

BDVe deliverable 2.1 of the work package on *Impact and Investment* is the first iteration of a living document aimed at documenting the makeup of an emerging European Big Data Value ecosystem. It *introduces the methodology* and provides a *first delineation of the ecosystem* as a high level *map* against which the impact of future actions and programs can be discussed and assessed. This first edition draws on recent studies and consultations with selected incumbents and new entrants such as start-ups and SMEs. It aims to inform policy discussion and action directed at enhancing the position and investment of European businesses and research in a globally competitive European Data Economy. The rolling report will include and inspire blogs and other short reads to be shared on the BDV portal *reaching out to and engaging* actors involved in the Public-Private Partnership and the wider Big Data Arena in Europe. The methodology includes (i) an approach to track Big Data developments and challenges across four dimensions: *policies, markets, technologies, and people*; (ii) in-depth analysis of selected topics of interest to the PPP; and (iii) preparation and sharing on the BDV portal and other channels of *thought-provoking short-reads and position papers*. The first two in-depth topics included in this report are treatments of *Business models for Data-driven Innovation (DDI)*, and *Free Flow of Data Policy*.

Takeaways from our review of DDI business models are that the added value within DDI does not necessarily lie in collecting mass data, but in finding valuable patterns and new forms of customisation and personalisation, placing the onus more on Artificial Intelligence and Machine Learning. With the expansion of Big Data - related techniques, also the amount and type of roles, types of activities expand (data retention, data marketplaces, data brokerage etc.) and in turn novel business models appear. Another finding is the beginning of a shift towards data driven innovation in verticals as well as combined horizontal and vertical models. The EU has been lagging behind in data marketplaces, but EU-based initiatives are emerging that want to offer EU-specific and EU-compliant databases. Finally, Data Science and Data Analytics are still a relatively small field of expertise; in order for vertical and combine business models to emerge and thrive, data scientist needs to be educated that have domain-specific knowledge on top of data science knowledge.

The special feature on *Free Flow of Data Policy* looks at the status and impact of the upcoming Free Flow of Data legislative proposals and the Public data sharing
D2.1: Report on high level consultation

initiative. It concludes among other things that the implementation may be complex and possibly slow given the broad impacts on current right holders.
1 Introduction

1.1 Scope of the deliverable

The objective of the first task of the work package on Impact and Investment (WP2) is to provide a continuously updated view on key issues related to the emerging structure and dynamics of the Big Data Ecosystem in Europe. It will serve as input to formulate concrete policy options and strategic actions to promote a seamless, competitive EU Data Ecosystem. It will inform other BDVe Tasks and Big Data PPP projects and engage other EU Big Data related initiatives on a strategic level. The task comprises the following activities:

- **Structural analysis, drivers and barriers**: insight from existing initiatives and materials will be compiled and complemented with consultation among stakeholders on perceived drivers and barriers to an effective European Data Economy.
- **Policy options and actions**: based on the input from the consultation, policy options and actions will be formulated that aim at addressing bottlenecks and that build on key strength of the European Ecosystem.
- **Discussion papers and action plans**: to share the insights of this work package, position papers will be prepared and shared through the BDVe network on selected topics emerging from the consultations and the policy analysis.

The current report presents the methodology and baseline for the structural analysis as the first iteration of a living document documenting the makeup of an emerging European Big Data Value ecosystem. It draws on existing and emerging studies, consultations with incumbents and new entrants such as start-ups and SMEs. It will identify emerging issues as input for policy action directed at enhancing the position of European businesses and research organisations in a globally competitive European Data Economy.

1.2 Document structure

This report should be read as a living document, meaning it reflects a snapshot of ongoing updates regarding BDVe and as such, is not a complete or rounded-off document. The deliverable is to provide a starting point and overview of what we set out to do, has been done and what key challenges for policy and business we can
D2.1: Report on high level consultation

already point out. The structure of the document is as follows: after the introductory section, chapter 2 will provide an overview of the methodology employed to arrive at the current deliverable as well as the approach to update the live document. Chapter 3 discusses the overall structure of the Big Data landscape and how to document it. Chapter 4 and 5 will consist of ‘special features’ respectively on business and policy. Chapter 6 will summarise preliminary findings and set out the main challenges and next steps for this WP.
2 Methodology

2.1 Overview

As a living document, this report aims to provide high-level, comprehensive and informative updates on the European data ecosystem. In this report, the methodology is to capture the rapidly changing landscape of Big Data ecosystems by consultation and condensation of a myriad of sources, including recent high-level reports, journal articles, EU policy briefs, blog posts, surveys and interviews with BVD-PPP stakeholders. Big Data and data-driven innovation ecosystems are rapidly evolving and diverging phenomena. Being potentially disruptive to current sectors and their accompanying business models, any attempt to capture an exhaustive list only once would not reflect the dynamics of this disruptive transformation. Instead, we propose to deliver periodic updates. The areas covered include policy, business and research updates on relevant aspects, fields or sub-domains of the European data ecosystem. The aim is to do so by looking into particular cases or sectors in which data-driven innovation is yet to take off, or is promising, but in full development.

More specifically, we will inquire into data value chains and ecosystems in sectors that could be considered as ICT-conservative or where the role DDI is hard to foresee, for different reasons, such as agriculture, transport or energy. Below, a set of steps and activities will be elaborated upon that together provide a methodology which will capture emerging, sometimes sector-specific trends and key issues that in turn contribute to the description of higher-level data value ecosystems and their accompanied value chains. The report will provide:

- A high-level-analysis of key issues and challenges within the BDV. From a top-down perspective, this analysis will show how and where novel BD and DDI trends are emerging.
- Mapping this analysis onto the status quo. A bottom-up periodical update on big data value chain developments and initiatives within the EU.
- Engagement. A series of short-read insights and provoking thought-pieces by key players in the field on big data ecosystems throughout the project.

Together, these three steps provide a way to capture oft-rapid developments within the BDVe, whilst also covering a variety of view. We will do so by taking the following steps:
D2.1: Report on high level consultation

A first baseline overview will be created by drawing on recent reports that provide high-level studies and analysis of Big Data Ecosystems. The scope of the literature search will be technological and/or on a policy level (meaning, relating to big data technologies and infrastructures, and the way they are guided and facilitated by policy). The focus will be mainly literature reflecting European developments, organisations, or treaties, possible augmented by sources external to the EU (such as OECD reports and key stakeholder white papers).

A second step is to periodically provide updates on changes and developments in policy, business or research deemed relevant for the consortium. Via literature reviews, document analysis and interviews with BDV-stakeholders, lighthouse-project leaders, sector-specific stakeholders and high-level report authors, we aim to capture trends and predictions, but also risks and fears stakeholders foresee in the Big Data ecosystem in the BDV-PPP at large, and within the sectors chosen by WP 2 specifically.

- For topical issues related to the ecosystem that render specific attention, we introduce special features. These are thematic focus pieces on one aspect of the ecosystem in which we provide an update on strategic insights or developments of relevance for the BDV-PPP. In this report, we offer one on business models (section 5) and one on EU regulation in relation to the free flow of data (section 6).

The third step is focused on choosing and evaluating the most effective form of verification and dissemination of our findings. We aim to do so by providing short reads in the form of: a) a series of blog posts for the BDV-PPP portal, either written by WP or in strategic co-authorships with key voices of a particular subfield and b) a set of position papers that will provide a more in-depth analysis of a particular issues or challenge connected to ecosystems and value chains of data.

Sectoral approach and lighthouse projects

To allow an in-depth discussion of the Big Data Value Ecosystem, work package 2 will zoom in on two new sectors each year. The focus sectors selected for the first year are Agri-food and Transport. Each of these sectors has a corresponding lighthouse projects. They cover sectors where the adoption of data-driven innovation is still relatively limited (Agri-food), and one that is a little bit further ahead (Transport).
2.2 Sources, data gathering and content development

Content of current and subsequent iterations of the report will be collected through desk study, interviews, and panel discussion at BDVe and other events. Feedback on the content will also be collected through regular strategy blogs on the BDV portal.

In the remainder of this section we will provide an overview of the information gathering and dissemination activities. It will be updated in 6-monthly intervals. Also, an overview is given of the partners we have aligned with to provide additional information and insights.

Per activity there will be short description of what has already taken place (M1-6) and what is scheduled to take place (M7-9).

**Blogs (ongoing)**

*M1-6*

Preparatory works for the first set of blogs to appear on the BVD-PPP portal, to be published from September onwards. The setup of the blogs is such that it aims to invoke discussion among BVD-PPP partners and beyond, also making use of experts as guest bloggers and/or co-authored pieces on a particular topic or case.

*M7-9*

A regular feature of the live consultation report are the strategy blogs. They intend to make key topics of interest accessible and engaging for the PPP actors and the wider BDV Ecosystem. The blogs will focus on emerging big data topics relevant for the main PPP audiences: (i) policymakers, (ii) companies and (iii) researchers. Extended versions are included in the report while shorter versions will be posted in a dedicated section on the BDVe Portal (WP5). The topics chosen for the first blogs include (i) Free Flow of Data and (ii) Big Data value propositions in Transport. The research for the Transport feature will also look at promising Big Data application scenarios to into the opportunity assessment of Task 2.2.

**Section on the BDV Portal (ongoing)**

A full portal covering Big Data Value PPP content and activities will be developed in Work package 5. The content created in Work package 2 on Investment and Impact...
D2.1: Report on high level consultation

aims to populate a dedicated section\(^1\). For this we have proposed a user-friendly, interactive design that offers a format and layout that allows easy reads and navigation. To illustrate this design we include a reference to a similar implementation on the pdpEcho website\(^2\) (Figure 1). For intuitive navigation we propose to use keywords that link to the source\(^3\).

---

\(^1\) The exact layout is currently being discussed as part of the work on the final portal design. The portal will likely come online in M7 or M8.

\(^2\) See pdp echo at http://pdpecho.com

D2.1: Report on high level consultation

strategic topics in these domains. The Free Flow of Data interviews and workshops also yielded potential co-authoring partnerships in this area⁴.

Content partnerships

The first achievement here includes an in principle agreement with leading Big Data blog ODBMS.ORG for them to conduct publish strategic interviews with key players in the BDV PPP. Interviews will be posted both on ODBMS and the BDV Portal.

M7-9

Interviews will be posted both on ODBMS and the BDV Portal.

Workshops

M1-6

For this report and the special feature on *Free Flow of Data contained in chapter 6* the BDVe team participated in the following three EC workshops on the topic of Data Access:

- **The Transformative Effect of Access and Reuse of Data for Smart Industries.**⁵ On 6 June 2017, we have attended the workshop on the transformative effect of access and reuse of data for smart industries. This workshop was organised by the European Commission and created an opportunity for BDVe to provide input on the challenges faced when trying to access and reuse data. Moreover, this workshop served as an opportunity to complete a list of consultations, studies and meetings on the “Building a European Data Economy” package launched by the European Commission earlier in 2017. Three main questions were raised and discussed during the workshop: (i) How is data transforming the European industry and society? (ii) How is data shared within the European data ecosystem? (iii) Which are the main challenges in accessing and reusing data? The key message was: (i) focus should be on what benefits farmers and (ii) industry is working to facilitate data exchange and sharing through trust and harmonisation.

- Workshop on *Data access and transfer with a focus on APIs and industrial data platforms*⁶ held on the 8th of June, 2017. The goal of the workshop was to explore the role of application programming interfaces in B2B (and B2C) data sharing. The main topics addressed that day were data sharing practices within Europe, the differences between public-and open data, the long tail and lock-in

---

⁴ Names of the companies/organisations concerned to be announced upon conclusion of the agreement.


models of hardware generated data in for example the automotive industry, sector-specific versus broad regulation in Europe when it comes to data sharing and modelling business based on data-locking models and the possibilities of APIs as a place where technical and legal rules can be shaped and documented.

- **Access for public bodies to privately-held data of public interest.** On 26 June 2017, we have attended the workshop access for public bodies to privately-held data of public interest. This workshop responded to the mid-term review of the Digital Single Market strategy which called for the need to further explore the issues of public bodies’ access to privately held data which are of public interest. This workshop was organised by the European Commission and created an opportunity for BDVe to provide input on the issues faced when trying to access and reuse data that are of public interest. The objective of this workshop was to discuss with stakeholders from public sector bodies and data holding companies the needs for access to privately held data and the conditions under which such access could possibly be granted. The three specific topics that have been discussed were: (i) access data for official statistics, (ii) access data for specific purposes in the healthcare sector, (iii) access data in the context of smart cities.

**M7-9**

For the upcoming European Big Data value Forum (EDVF) in November 2017 in Versailles, WP2 proposed a session on ‘European Big Data Ecosystem: Investment and Impact’. It aims to discuss key strategies to drive industrial impact and investment as well as review progress on addressing the constraints on data-driven innovation in the EU. The second part of the session will dive into the dynamics and levers of the underlying innovation ecosystems. On the basis of concrete examples and best practices, big data business opportunities will be discussed. The session will conclude with a discussion of approaches towards an engaging BDVA marketplace to promote the data-driven innovations of the PPP to a wide audience of Start-ups, SMEs and Investors.

**Interviews**

For this report and the special feature on *Free Flow of Data contained in chapter 6* the BDVe team has conducted interviews. The list of interviews and the list of the questions can be found in the Appendix A, B and C. The conduction of interviews will be an ongoing process.

---

3 An Emerging Big Data Ecosystem in Europe

3.1 Introduction & Methodology

This chapter aims to provide a first snapshot of the emerging European Big Data Ecosystem. To do this, we follow an approach presented by Lawrence Lessig (1999) in his influential and comprehensive publication ‘Code and Other Laws of Cyberspace’\(^8\). Lessig suggests that online and offline enabling environment (or ecosystem) as the resultant of four interdependent, regulatory forces: Law, Markets, Architecture, and Norms (See Figure 2 below). He uses it to compare how regulation works in the real world versus the online world, in discussing the regulability of digital worlds or cyberspace as it was called in 1999.

![Figure 2: Four regulatory forces (Lessing, 1999)](image)

In our report we propose to map the development of the Big Data Value Ecosystem as the sum total of developments along these four dimensions. The Strategic Research and Investment Agenda (SRIA) of the BDVA follows a rather similar subdivisions in Technical (Architecture) and Non-Technical Priorities: Business (Markets), Policy (Legal), and Social (Norms). This ensures that our findings can inform progress on the priorities of the BDVA and the BDV PPP. In the analysis we

---

D2.1: Report on high level consultation

equate Lessig’s notion of ‘Environment’ to the BDVe concept of ‘Ecosystem’. Before we start our mapping we need to review briefly what we actually mean by ecosystem and how it relates to similar concepts also common in the language of the EU such as Clusters and Hubs.

3.2 On Ecosystems, Clusters and Hubs

The notion of (business) ecosystem, widely spread and commonly used nowadays in business and government circles alike, comes from a theoretical point of view rather close to Porter’s cluster concept, especially when the term ‘associated institutions’ is interpreted broadly. The original cluster was introduced in Porter (2003) as “geographically close groups of interconnected companies and associated institutions in a particular field”, linked by common technologies and skills. They normally exist within a geographic area where ease of communication, logistics and personal interaction is possible. Clusters are normally concentrated in regions and sometimes a single town”. Porter’s cluster concept strongly builds on the notion of agglomeration effects (agglomeration economies) long recognised in economic geography and regional science literature, dating back as far as 1890 to the work of the English economist Alfred Marshall. A similar recent line of thinking can be found in what is called the new economic geography, which started with the work of Paul Krugman in Geography and Trade (1991).

The idea of business ecosystem stems from a clear analogy of the business world, its players and dynamics with biological ecosystems defined as a biological community of interacting organisms together with their physical environment in which capabilities and roles co-evolve (Moore, 1996). Similarly, a business ecosystem can be defined as a network of buyers, suppliers, producers and providers of related products or services set within a certain institutional and regulatory context. Moore (1996:26) originally defined a business ecosystem as “(a)n economic community supported by a foundation of interacting organisations and individuals—the organisms of the business world. The economic community produces goods and services of value to customers, who are themselves members of the ecosystem. The member organisms also include suppliers, lead producers, competitors, and other stakeholders. Over time, they coevolve their capabilities and roles, and tend to align themselves with the directions set by one or more central companies. Those companies holding leadership roles may change over time, but the function of ecosystem leader is valued by the community because it enables members to move
D2.1: Report on high level consultation

toward shared visions to align their investments, and to find mutually supportive roles.”

There is not one single archetype business ecosystem - different forms may occur, co-exist and evolve over time. Along with the notion of business ecosystems, recently an increasing interest in innovation ecosystems can be observed (see e.g. Adner and Kapoor, 2010). A common finding is that organisations form an ecosystem around a common core (Talvitie, 2011), i.e. platforms, technologies, processes, standards or other assets common to and used by members of the ecosystem in their businesses. What should be clear also is that an ecosystem is an abstraction that can be applied at different levels, from an entire industry, a small consortium of companies to also a defined subset such as a digital ecosystem (e.g. Nachira et al., 2007). The evolution of a start-up ecosystem should be seen from this perspective. The European Big Data Value Ecosystem is formed in a similar vein, connecting Data and ICT providers with Data service providers including start-ups, and horizontal and vertical End-users and consumers.

At the extremes, two business ecosystem ‘models’ stand out, the so-called keystone model as originally implied by Moore (1996), dominated by a large firm (the keystone) interacting with a large number of small suppliers (Iansiti and Levien, 2004). The keystone model would be typical for the US. Another model, the so-called flat model would be more typical for Europe, composed of mainly small and medium firms, but accommodating also large ones (e.g. Corallo et al., 2007). Put differently, business ecosystems can generally have either a hub-centred star structure or a flat mesh-like structure. The presence of hubs would make the network robust to the removal of individual nodes, that is as long as the hubs are there (Iansiti and Levien, 2004; see also Mazhelis et al., 2012). A hub is used here as in quantitative network analysis, as a component of a network, a high - degree node, i.e. one that has significantly bigger number of links in comparison with other nodes in the network. The EU initiative on Digital Innovation Hubs is an example of a hub - based approach. The BDVA iSpaces could also qualify as a Digital Innovation Hub.

The term hub, however, is also used in a more general way, to delineate “places with high levels of company and/or employment specialisation” (Mateos-Garcia, 2014). The term hub can also refer to a specific type of cluster, so-called hub and spoke clusters, dominated by one or several large firms surrounded by smaller suppliers and related activities, buying from or selling to an anchor firm or to take advantage of activities attributed to the anchor firm’s presence (OECD, 2007; Barkley and
Henry, 2001). Note the similarity, again, with the preceding business ecosystem typology.

Within the context of this study *innovation clusters* are understood as “the geographic entities that encompass start-ups, accelerators, incubators, investors, research centres, smart communities, etc., operating in one of the IoT areas (like, e.g., smart living environment, smart farming and food security, wearables, smart cities, smart mobility, smart water management, smart manufacturing).” The OECD proposed a mapping of the actors in the global data ecosystem (Figure 3 below).

![Figure 3: Mapping of actors in the global Big data Ecosystem](image)

In upcoming editions of this report we will document the structure of the EU data-ecosystems by mapping key actors and roles on the value chain and by keeping track of emerging big data hubs and clusters such as the iSpaces and the Digital Innovation Hubs. In the next section we will explore the main four dimensions creating the enabling framework of this Big Data Landscape.

### 3.3 Law and Policy

The October 2013 European Council concluded that: “EU action should provide the *right framework conditions* for a single market for big data”. The Commission responded with an outline for *a new strategy on data-driven innovation, supporting and accelerating the transition towards a data-driven economy in Europe.*
D2.1: Report on high level consultation

In order to boost investment and foster community building the Big Data Value PPP, a contractual Public-Private Partnership on data was created in October 2014 with the objective to fund “game-changing” data innovation ideas, building on a Strategic Research and Innovation Agenda (SRIA). Close to 180 members of European Industry are represented in the Big Data Value Association that accompanies the PPP. The current BDVe project was funded by the commission to support the wider goals of the PPP on increased investment, impact, community-building, and skills development in the area of Big Data. Total Big Data Investments into Research and Innovation as a result of the PPP are expected to reach around EUR 2.5 billion over 2016-2020 period.

As part of the Digital Single Market (DSM) strategy, the Commission recently adopted a Communication on “Building a European Data Economy” aimed at fostering the best possible use of the potential of data to benefit European economy and society. Beyond the protection of personal data, it covers issues such as access to machine-generated data, protection of confidential data and lock-in effects. Finally, for the end of this year and the first half of 2018 the Commission is preparing a legislative proposal on a new EU Free Flow of Data cooperation framework as well as an initiative on accessibility and reuse of public and publicly funded data. It will addresses key framework conditions such as interoperability, usability and access to data in business-to-business, business to consumer, machine generated and machine-to-machine data. For a detailed discussion on these latest proposals see our policy special feature contained in chapter 6.

3.4 Market

The most recent update of the market analysis of Big Data is provided in the European Data Market Monitoring tool and other recent IDS-reports (see Figure 4 on the next page).
These estimates include all the direct and indirect economic impacts produced by the adoption of data-driven innovation and data technologies in the EU. The composition of data companies includes start-ups, innovative SMEs, and existing enterprises that exploit the emerging business opportunities of the data market. Data companies’ are registering a considerable growth rate of about 7% year-on-year. Although the data sector is thriving, a large scale uptake by data users in verticals is still lacking.

*In our business special (chapter five) we will discuss recent developments in DDI business models for the EU Big Data market.*

### 3.5 Architecture and Infrastructure

Architecture and Norms will be discussed in more detail in future editions of the report. This section includes a brief consideration of *technologies and infrastructures* enabling the European Big Data Ecosystem. The SRIA covers the technological priorities already in quite some detail. In upcoming editions of this report we will focus on recent developments towards a common EU reference architecture, standardisation and any emerging technology scenarios that may inform the SRIA.
D2.1: Report on high level consultation

The most recent development is an updated high level technical architecture schema prepared in the BDVA Technical task force. It is currently being discussed with the lighthouse projects, Big Data initiatives, as well as related initiatives in the domain of IoT and HPC.

![Figure 5: Draft Big Data Reference Architecture](Source BDVA AG, May 2017)

Relevant topics that will be considered for discussion in our next report are:

- **Alignment of architecture developments.** There are multiple initiatives across Europe developing reference models with (Big) data components. These include for example communities working on HPC, IoT, Future Internet and Industry 4.0.
- **Standardisation and Interoperability.** It is argued that Europe is not sufficiently active in the main standard setting bodies. A BDVA taskforce is now dedicated to address this. What are the best strategies?
- **EU Open Data Portal⁹.** We will be looking into developments on European Commission initiatives on supporting and fostering open data in the EU.

---

⁹ The Commission launched a fully-fledged pan-European infrastructure to federate content published on European public open data portals. The portal now provides access to more than 400,000 datasets from all over Europe.
3.6 Norms

Understanding the social dimension is essential to ensure that DDI creates value in a way that respects and promotes personal and collective wellbeing. Although the use of new data sources and Big Data methodologies is expected to have a positive effect on policy making, resulting in evidence based and better tailored policies there is a clear risk of negative outcomes such as enforcing inequality and discrimination. A recent WIRED article describes the complexity of reviewing the ethical implications of Big Data applications well by pointing to high profile cases where massive public outcry caught review boards completely unaawares.\(^{10}\) In their whitepaper, the Council for Big Data, Ethics, and Society (BDES) \(^{11}\) finds that:

“there is a disjunction between the familiar concepts and infrastructures of science and engineering, on the one hand, and the epistemic, social, and ethical dynamics of big data research and practice, on the other”.

The paper argues that to advance ethical conduct in data science “requires careful consideration of big data’s broad technical, social, and political contexts”. It goes on to suggest new policies that promote greater reflection and engagement. The first tranche of PPP projects include four projects on privacy-preserving big data technologies and one on Enabling responsible ICT-related research and innovation. WP2 will engage with and support these projects to define common topics and to ensure key findings are included in the broader narrative of the BDV PPP.

3.7 Sectorial Perspective

To allow an in-depth discussion of the combined effect of market forces, architecture, policy, and norms on the creation of Big Data Value Ecosystem, work package 2 will zoom in on two new sectors each year. The focus sectors selected for the first year are Agri-food and Transport. They have been selected for two main reasons: (1) Each of these sectors have corresponding lighthouse projects, and (2) they both cover a sector where the adoption of data-driven innovation is still limited (Agri-food), and one that is a little bit further ahead (Transport). For the BDVe it is

\(^{10}\) [https://www.wired.com/2016/05/scientists-just-confused-ethics-big-data-research/](https://www.wired.com/2016/05/scientists-just-confused-ethics-big-data-research/).

\(^{11}\) The BDES brings together researchers from diverse fields involved in the ethical, social and policy challenges associated with the rise of “big data” research and industry. See [http://bdes.datasociety.net/wp-content/uploads/2016/05/Perspectives-on-Big-Data.pdf](http://bdes.datasociety.net/wp-content/uploads/2016/05/Perspectives-on-Big-Data.pdf).
D2.1: Report on high level consultation

more useful to develop insights into sectors. Both sectors are introduced briefly below.

*Transport, Mobility and Logistics*

Big Data will have a profound economic and societal impact on mobility and logistics according to the BDVA taskforce on transport. The sector contributes to 15% of GDP, employing 5% of the workforce in the EU. The OECD calculates 500 billion USD in potential value from operational efficiency gained with Big Data. The sector is expected to grow rapidly while generating and managing increasing amounts of data\(^\text{12}\). While there is an increased interest in the use of Big Data applications, the uptake is still limited with only 1 in 5 companies employing Big Data solutions.\(^\text{13}\) As a key contributor to CO2 emissions, efficiency gains in the transport sector result in substantial reductions to greenhouse gas emissions. With many companies in the sector quite familiar with data collection and processing tools and the relatively limited legal constraints on the wider use of logistical data, there is great potential for extensive use of Big Data technologies and applications in Europe.

*Agriculture*

According to estimations from the United Nations, over ten billion people will inhabit the planet by 2050. Simultaneously, climate change is diminishing arable land and resources. These developments put pressure on the agriculture and food supply chains, which need to evolve and adapt current practices in order to dramatically increase their output to meet (and cope with) a fast growing demand for food while improving efficiency in terms of the use of water resources, energy and materials. Parallel to the challenge to feed the world’s population, there is an increasing consumer demand for safe, high quality food, which is produced in a sustainable and transparent value chain. This shift in attitude of consumers, in combination with regulations regarding food safety and sustainability, force the retailers and the food processing industry to respond. They need to be transparent in the origin of their products, how these goods are being produced and whether they are indeed safe and produced in a sustainable fashion. Kruize et al. note that these developments will have a big impact on food production operations, management and information exchange between stakeholders involved “[...] because production agriculture always has performed and always will perform essential tasks in the web of the food industry.” New technological innovations are essential to address these challenges.

\(^\text{12}\) Source Alice ETP.

\(^\text{13}\) Source PWC.
Data-driven innovations in particular - i.e. precision farming, cyber-physical systems, advanced business management software, monitoring technologies and genetic enhancements - can be a driving force. Furthermore, as agricultural enterprises are increasingly becoming bigger the need for both automation of physical tasks (i.e. milking) and information management becomes more urgent, which also points to a more central role for data. In chapter 5 and 6 we will return to discuss DDI models in agriculture and transport.
4 Business Special – Data Application Scenarios

4.1 Introduction

This first special feature section will describe current understanding and perspectives on generic business models and application scenarios for Big Data. In the first edition of it will focus on setting the scene and reviewing methodology as well as highlighting some of the application scenarios in the year one-focus sectors Agri-food and Transport.

4.2 The (general) business case for big data

One of the many promises attached to data-driven innovation, in particular with Big Data, is the potential to disrupt business and create novel chains that can potentially reshape entire sectors within an economy. The main line of reasoning in a recent OECD report is that due to the vast and rapid proliferation of computers, sensors and actuators in virtually all sectors of society, we can witness a process of mass datafication of business processes and government services. ICTs and Big Data have the potential to connect sectors, redefining their value chains. Recent reports and papers emphasise that there is no single business model for big data. However, some defining characteristics and conditions can be identified, as well as some of main risks and pitfalls.

In this report, and in the project as it continues, we aim to capture some of the key trends in Big Data business models as well as their application in key sectors of the European Economy. We will focus on data-driven business models (i.e. what works and what not), yet with the caveat that they are difficult to define and scope. First of all, the data market is evolving fast, making it difficult to see which business models will ‘stick’. Secondly, as this market is driven by networks that are organisationally and technically globalising, consequently also a variety of cultures and notions of business models are mixing. Although, particularly in the B2C market, the data market is still very U.S.-driven with many large ICT-providers based in the US, other continents, countries and also other business cultures and models are entering the

---

14 OECD report on Big Data Ecosystems.
15 See footnote 23 and NESSI whitepaper.
fray. The challenge for the BDVe project is to isolate developments in data-driven innovation and value creation that are relevant and important in a European Digital Single Market.

4.3 Big Data Value chains and ecosystems

(i) What is in a data value chain?
As described in the SRIA of the BVD-PPP\textsuperscript{16}, the approach to be taken to research data-driven innovation is to take the perspective of a process-driven approach (rather than a rule-based or statistical-based approach)\textsuperscript{17}, by which it is meant that in order to capture added value in Big Data or Data Driven Innovations, the approach is to capture the processes of (sub) parts of along a value chain. Among others, Cuddington et al argue that a process-approach is best suited when researching ecosystems on a (semi) global scale. In our case, the first level of analysis would be to identify or roughly sketch out the contours of a data value ecosystem. In general, the data value ecosystem comprises of a set of data value chains. These value chains are generally described as consisting of the following key elements (see also figure 5 in chapter 3 on a recent visualisation of the BD value chain):

- data generation or acquisition
- data analysis and/or processing,
- data storage and curation, and
- data visualisation and services (the so-called the last mile of data driven innovation).

In each of these four steps, added value can be created and although all steps of this value chain can be present within one company, they do not need to be. Data driven innovation can take place within each one of these four steps or in combinations of these steps. Equally so, while DDI is consistent of these steps, the increase of complexity within DDI also means that having all four steps in-house, or developing them in-house less likely to be a fruitful route. Consequently, DDI relies more on interdependencies and mutual risks, and if both the amount of data and the types of analysis are growing, there will be an increasing need of strategic collaboration. When delving into further detail of those four steps, some key activities are linked to each of the steps. As depicted by the BVDA\textsuperscript{18}, the value chain follows the following steps within the big data ecosystem.

\textsuperscript{16}http://www.bdva.eu/?q=sria.
\textsuperscript{17}See Cuddington et al. 2013.
\textsuperscript{18}Diagram taken from BVDA SRIA report version 3, p. 24.
D2.1: Report on high level consultation

(i) Potential added value can be categorized in what is being offered to a customer (be it B2B or B2C). One can imagine a startup that has to buy raw data from another party already has a different take on how to create added value than a company that is generating data themselves and ‘only’ has to invest in an analytics department. Moreover, one can differentiate between companies who generate data as a byproduct of their core activities or companies whose core activity is to generate and store data. The newness, and major potential for disruption of data-driven business models is that they can shift a company’s core activity or turn it completely upside down: from the famous example of John Deere shifting from being a tractor company to a provider and analyser of information such as maintenance scheduling\(^\text{19}\), to more recent examples of Tesla starting to offer in-car entertainment\(^\text{20}\), entering a competition with existing media streaming platforms.

(ii) These are examples of a specific sectors (respectively agriculture and transport) utilising ICT and Big Data to alter their business model entirely. However, most Data Driven Innovations (DDIs) stem from the ICT sector itself. In this sector, we can see that in star-structured hubs as described in section 3.2, some nodes become too big and powerful, leaving little room for smaller nodes to flourish. Here we can witness a move towards a ‘winner-takes-all’ strategy, in which companies or organisations within the ICT sector try and experiment with data-driven innovation models in

\[^{19}\text{See Cavanillas et al. 2016.}\]
sectors they were not previously active in whatsoever, causing disruption and potential monopolisation (see f.i. the big five of GAFA\textsuperscript{21}, also dubbed Over-The-Top, or OTT media companies\textsuperscript{22}). Such companies often enter by offering data analytics, to then also start offering other data-related services to the non-ICT sector client. They can for instance equip a factory with sensors, and in the process also take over data storage as a service and potentially as a source of a new business opportunity. In order to offer services on all the 4 steps as depicted in figure 6, a DDI actor needs to have the infrastructure, facilities, expertise, financial room, and most importantly, data, to experiment and explore new business models - this leads to large players often being the only ones able to take on such roles.

(iii) Many start-ups and more mature data-dependent companies are relying on large data warehouses, meaning they are reliant on buying or otherwise making use of external data via for instance subscription models, simply because the start-up costs of developing and maintain one’s own data warehouse are far too high. Business models might be moving away, or evolving on top of, raw data creation and storage towards creating added value through analytics and insights, and to making such insights available in the right form for the right customer. Or, to put it differently, the problem for Big Data is finding small patterns\textsuperscript{23} and being able to offer those puts one in a competitive advantage.

(ii) Where can added value be found in DDI?
According to Erbes et al, data driven companies and ecosystems can become successful if they (i) pay attention to data flows as opposed to stocks, (ii) if they rely on data scientists and product and process developers rather than data analysts and (iii) if they are moving analytics away from IT function and into core business, operational and production functions\textsuperscript{24}. These three directions clearly indicate that companies, if aiming to innovate through use of data, need to take the role and the place of data within the business model far more serious. In order to get a grip on different forms and types of added value that can be obtained through data, Hartmann et al. (2014), after having analysed a wide range of papers on data driven


\textsuperscript{24} See Erbes et al. 2012.
business models and value chains, conclude that there are three discernible types of offerings:

- raw data;
- information or knowledge;
- non-data or non-virtual offering.

Further delving into how these three offerings would translate into different business models; they provide a data-driven-business-model framework (Hartmann et al. 2014) that helps in grasping the different possibilities and flavours of business models currently available in what they call a data-driven business model (DDBM). In this model\(^{25}\), a differentiation is made in terms of what can be done with data on different levels, also mapping out further steps within this differentiation. They discuss the data source (f.i. internal or external), the key activity planned with the data (f.i. data generation, processing, analytics etc.), the type of offering as mentioned above, the target customer (B2B or B2C), the type of revenue model that is aimed for (licensing, freemium, advertising and so on) and finally the specific cost advantage. In terms of mapping which business models are currently out there, a recent OECD report\(^{26}\) on mapping the big data ecosystem also provides a comprehensive overview of models currently seen in both online data-driven environments and actors such as GAFA and other industries and sectors moving to data-driven innovation. They list the following models, which they have identified as most common, and the potentially accompanied actors in the Big Data Ecosystem (table 1):

<table>
<thead>
<tr>
<th>Model</th>
<th>B2B/B2C</th>
<th>Actors involved</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freemium</td>
<td>B2C</td>
<td>Users, content- and software producers</td>
</tr>
<tr>
<td>Advertisement</td>
<td>B2C</td>
<td>Ad companies/search engines (GAFA), multi-sides market hosts</td>
</tr>
<tr>
<td>Subscription</td>
<td>B2B</td>
<td>Media providers, ISPs</td>
</tr>
<tr>
<td>Usage fees / pay-as-you-go</td>
<td>B2B</td>
<td>Start-ups, XaaS providers</td>
</tr>
<tr>
<td>Selling of goods / per-per-download</td>
<td>B2B &amp; B2C</td>
<td>IT infrastructure providers, service platform providers</td>
</tr>
</tbody>
</table>

\(^{25}\) See Hartmann et al. 2014.

Selling of services | B2B | IT consultancy services, software development/maintenance / support; intermediaries
---|---|---
Licensing / IPR | B2B & B2C | Not widely used in BD ecosystem start-ups
Commission fees | B2C | Data analytics intermediaries

Table 2: Overview of business model
(Source OECD report)

As pointed out by this report and based on an analysis of a 100 start-ups performed by Hartmann et al. (2014)\(^{27}\), the choice of licensing is not very popular among start-ups. Moreover, they still maintain divisions between B2B and B2C based models, yet increasingly data-driven revenue models cut across these traditional divisions, due to a dependency of data that is being generated in different forms and constellations - ranging from machine2machine data and data generated between humans and machines - and/or a mix of those - between end-user and provided and between data buyer and data seller. One model that is not yet mentioned in the B2C or the B2B classification of models is where an end-user or company pays with data. This *pay-with-your-data principle* can work in different ways, for instance potentially the devices or goods that are offered are for free or offered at a lower price, in exchange for data collection and storage. Or, with the emergence of data markets, there might be forms in which data is not offered at a price, but at a demand for other types of data). Connected to such emerging business models between companies or between companies and consumers, forms of connected - *hardware based lock-in models* also emerge. One can think of a device that can only be used - or its added functionalities can only be unlocked if data is being shared (*freemium* models in hardware), or models where third-party services might emerge from data generated via hardware use. The *fitbit* is an example of an external third-party hardware sensor that adds data to a smartphone app - where rudimentary functionality works directly; supplementary data analysis is only available via a premium package\(^{28}\). Another example would be a smart doll that only unlocks higher forms of AI if the user allows for more data to be collected\(^{29}\).

\(^{27}\) Big Data for Big Business? A Taxonomy of Data-driven Business Models used by Start-up Firms - working paper by Philipp Max Hartmann, Mohamed Zaki, Niels Feldmann and Andy Neely - Cambridge Service Alliance 2014.

\(^{28}\) See f.i. https://revenuesandprofits.com/how-fitbit-makes-money/.

(iii) Who is doing DDI and why?

Besides the unpredictable development of new types of business models within DDI, the goals of DDI business models show less variability. Roughly two main reasons can be found to date for engaging with data-driven business models, which are:

(i) To improve or augment current business models in terms of performance enhancement due to more possibilities to do fine grained analysis on either parts of - or the entire business process.
(ii) To use data to alter a business model and strategy entirely (meaning to alter the core business due to insights of own data or combination internal and external of data).

In visual form, the BVDA SRIA report\(^{30}\) shows the following steps of where added value can be found:

![Figure 7: Steps in the data value chain](Source BDVA SRIA report)

Where actors in earlier-described vertical sectors would generally be more prone to ‘play it safe’, starting by augmenting current business models and values chains (the most right box of figure 7) the horizontal (ICT) sector would be more inclined to look for disruption (the most left box in figure 7). However, such divisions are increasingly difficult to make and combinations can be made of the above-mentioned reasons and accompanied actors. One way of getting a grip on the EU Big Data Ecosystem is therefore a careful selection of cases that can reveal how the horizontal-and vertical sectors interact when trying to ‘do’ DDI.

The contours of a Big Data value Ecosystem consists have been provided above, yet some key questions and challenges emerge that beg further research and exploration. For those sectors within the EU, and for the entire EU big-data economy, strategic challenges such as where and how to find partners from the

\(^{30}\) BVDA SRIA report, p. 37.
D2.1: Report on high level consultation

horizontal sector, how and where does disruption take place, when to jump onboard, where and how data can be stored, bought and traded, and how to be certain of its safety, quality and compliance to EU regulation are of key importance and this WP aims to track specific developments in upcoming editions of this report.

Figure 8: Model for the Big data Ecosystem
(Source TNO 2017)

In Big Data ecosystem models such as figure 8, we can see both the horizontal sector (ICT) and 2 sample vertical sectors, in this case agriculture and transport. If we assume that within Europe, the majority of sectors and businesses within that sector are somehow dealing with - or dependent on - computing, then we can also assume that there will be forms of data generation within each sector, leading to the possibility to a larger or lesser extent, of added value through data. The overlap of the horizontal with the vertical sector can be seen as irreversible and increasingly entangled in terms of business dependency. Think for instance of any business, (be it a farm or a logistics centre), where data would be generated on a daily basis in administration and accounting, communication, marketing, planning, vehicle...
performance, process optimisation etc. As can be seen in the model, this leads to different types of databases within a sector or business with potential new added value.

A separation is made between internal data, meaning data generated in, by or through the company itself and external data, meaning data that comes from outside the business or sector and is brought in. The division of internal versus external data can be further split via a technological axis - is data natively digital, such as sensor data, or is it digitised and thus ‘copied’ such as Google books - or via an organizational, strategic axis - is data being used as part of the core a business model or is it a by-product. The former axis leads to questions of regulation, IPR, data ownership, but also of technical formats and standards, and issues of data translations and decontextualization. The latter axis introduces questions of added value in relation to ICT infrastructure, investments and strategic dependency among others. If data-driven business models are not part of the core business, should they be, and if so, which data by-product seems promising and what investment is needed to actually create added value? Once the step is taken to start involving data in the business strategy, a next question would be if a company does this by itself, or if they would need strategic alliances with other actors from the horizontal and/or vertical sector? These and more related business strategic questions and challenges are especially relevant for sectors that have not yet engaged with DDI, or just got started, such as perhaps some parts of the agri-food sector. Hartmann et al. provides an insightful taxonomy of types of business strategy and connected data activities that aim to guide such questions via a categorization between data sources and data activities. They define 6 types of business models in figure 9 below:

---

31 The figure stems from “Big Data for Big Business? A Taxonomy of Data-driven Business Models used by Start-up Firms” - working paper by Philipp Max Hartmann, Mohamed Zaki, Niels Feldmann and Andy Neely - Cambridge Service Alliance 2014
D2.1: Report on high level consultation

Figure 9: Types of business models
(Source Hartmann et al. 2014)

The figure sets out the different types of data sources against the type of data processing activity. Earlier mentioned large players in the ICT sector would be type F, whereas start-ups and scale-ups would show (combinations of) a, b, d, e and f or fit into the profile c. These types can help us in mapping cases we want to research in this WP, and can serve as a guide to understanding how and where added value is being created, or sought.

Besides models of where and how to find added value in DDI, we consider it of high importance to mirror-and draw from-DDI practices in the current EU Big Data practices, Below a preliminary list is provided of key challenges and risks (or drives and barriers), based on interview workshops attended and recent literature.

4.3.1 Some key trends and challenges within DDI business models
Following the methodology as described in section 2 of this report, besides a high-level overview, we aim to capture more bottom-up lessons from DDI practices within the EU that are more closer to practical yet strategic points to take into account In follow-up versions of this report, we will mirror more in-depth the theory on DDI business models with case studies to draw valuable and implementable insights and recommendations. A preliminary list of trends and challenges:
The emergence of multi-sided markets within the EU. Data markets are places where data providers, brokers and buyers come together. Although mainly US-dominated, there are growing initiatives for data marketplaces within Europe.

Business model disruptions across the ecosystem: they happen both horizontally and vertically. This means innovation and added value need not come from within a sector, and although the most risky of all DDI variants, moving cross-sector seems to offer the most potential way forward.

From data-actors imposing on existing sectors to sectors becoming disrupters (through ICTs). Instead of relying solely on ICT sector - partners, literate suggests a move towards an increasing role for in-house data science or otherwise new horizontal connections through data.

PaaS, DaaS DIaaS - new types of services and actors emerge from these different forms of disruption. New forms of services emerge with new needs for (local) data storage, brokerage, access and other data-related activities that can turn into full-scale service industries.

The move from startups to scale-ups. Besides, or next to the startup as a way to commence DDI, the scale-up is gaining in visibility, the latter meaning a company that has a better and more mature idea on the product and how to monetise it and as such are seen as more stable and reliable potential partners for DDI.

APIs (application programming interfaces) as a new core piece of software and potential regulating mechanism. APIs form a crucial element in connecting different streams of software and data and are shaping up to become a very important category or of software and a crucial building block for DDI, also because within the API all sorts of access and property rights can be arranged (it can be seen - and built - as a contract).

From access-based models to databases, to action-based models to databases. In order to open up new possibilities for accessing data, alternative models form the all-or nothing access to databases start emerging, of which an ‘action-based’ models seems most promising (action-based meaning a company or actor is allowed to perform analysis technique X but not Y and to the granularity of Z, rather than access, based, in which company A has access for the price of B per month).

Key challenges for DDI business models:

Control points within the ecosystem can lead to key data holder - monopolies. When vital data-control points such as portals or data storage

---

32 See f.i. www.dawex.com
33 Respectively Platform, Data and Data Infrastructure as-a-service
34 See f.i. Isenberg, 2012.
35 API stands for Application Programming Interface.
cloud facilities are in the hands of a few, this can lead to monopolies and unfair competition.\(^{36}\)

- Data lock-in models and/or walled gardens in data platforms for businesses and consumers. This is both a risk and an opportunity - easy and streamlined connected services do offer added value for the end-user, yet, it runs the risk of becoming too dependent on the data-service chain, making it harder to switch or repurpose data.\(^{37}\)
- Lack of compatibility and challenges of standards in datasets. Connected to the former point, there is wide variety of types of programming languages, sensors and ways of processing and presenting data-driven insights. Standard operating procedures and ISO standards need to be developed in order to harmonise the DDI market.
- Lifespan of data and bottlenecks in information flows. Besides the earlier-mentioned control points that can be put in place deliberately, another hindrance for DDI is the lack of documentation standard and consistency, and the lifespan of data - when and for how long is data of any value and for whom and if so, how is that data available or how can a company make it available?
- Strategic choices on the data market. Connected to the former challenge, strategic key decisions to be taken are to choose between harvesting and collecting data and/or sharing or buying and selling it. With a rise in segmentation of DDI-related activities and tasks, there might be room for diversification of types of data markets, and platforms, which offers both opportunities and risks.
- Data analytics is accessible for experts only. Also of strategic nature is the choice for a company or organisation to invest in in-house data science, or to outsource this section, or to partner up (both with the risk of opening up). For the moment, data analytics remains an expert-based market.
- Free flow of data vs more strict privacy and data protection regulation in the EU. With the coming into force of the GDPR and other connected upcoming data-related regulations, the EU data market is gearing up to become a unique data place, which mean more expertise is needed for companies to deal with these regulations, while on the other hand it also offers privacy and data security as USPs.

\(^{36}\) There is regulatory pushback against monopolies in the Big Data Ecosystem, for instance the recent ruling in which a dominant player misused its position (https://www.theguardian.com/technology/2017/jun/16/google-fine-eu-case-shopping-search-result), however, many DDI business models are geared to monopolies, as often the revenue model only works or improves when the amount of users or data-points increase, leading to the development of dominant players (such as it has happened in the search engine- and browser wars, in which there is only room for a few or even 1 player).

\(^{37}\) This links to data portability rights for citizens and customers within the EU, yet it also refers to moving and opening up vs locking in data streams in a b2b context.
Summarising the above, a variety of mixed mode horizontal/vertical DDI models can be distinguished at sectors and business level, each with many challenges on both a technical and an organisational level. As a consequence, being able to pinpoint and predict where and how added value will emerge is increasingly difficult if not impossible with current approaches. However, by adopting a sector-specific approach as the starting point, we aim to find important pointers and directions for DDI application scenarios that can say something about how DDI-low sectors might evolve. From that we aim to derive more general insights into the EU Big Data Ecosystem.

The next section will contain a brief discussion and description of the approach we aim to adopt, to be followed by a section in which two short scenarios will be sketched on the consecutive sectors of agriculture and transport.

### 4.4 Two short scenarios: Agriculture and Transport

Following the methodology as described in section 2, the approach taken to try and uncover different aspects of the EU Big Data value ecosystem is to investigate specific sectors. In support of the wider BDV PPP the BDVe project aims to help identify areas with untapped DDI potential in Europe while also learning from sectors that are a little bit further ahead. Hence the choice for the DDI low Agri-food sector and the Transport sector where there has been some significant development. Based on a first set of preliminary interviews, a brief overview of business model scenarios will be provided for respectively the agricultural - and the transportation sector.

#### 4.4.1 Making crops smart - potential DDI in Agriculture

Agriculture is inherently linked to technological change and novel techniques – not only on the level of how to cultivate the land and to steer how, where and which crops or livestock can be developed and maintained. However the use of ICT and Data is still rather limited. From a larger governance-perspective the question of how to predict the amount and type of production from the land has been crucial and instrumental in the development of the use of statistics in modern Western-European modes of governance. In this regard, to consider agriculture as a data-low and conservative sector is not entirely doing it justice, historically speaking. When taking a more recent perspective, some movements can be witnessed that range from up-scaling to large scale mega-farms to the return of local, small bio- and
organic ways of farming that specialise and tailor to specific market demand. These seemingly diverging directions of business models are in flux and not necessarily mutually exclusive. One additional aspect that is not yet very much on the radar in capturing, guiding or instigating novel directions in agriculture, is DDI. If we return to the model as presented above, several opportunities can be imagined for DDI in agriculture that poses different strategic and policy-related challenges. If we were to apply the key lessons of data driven business models to agriculture, the following directions emerge (not aiming to be exhaustive):

**Data to improve internal business models**

With the intensification of agriculture, meaning an increase of production per piece of land and a decrease of personnel, also the need to monitor and control crops or stock has grown. There are many innovations in the agricultural sector that support and drive this intensification. Such innovation can be found in automated feeding-machines for livestock or localised crop management based on periodic drone footage. In general, such innovations are ICT-based and have a strong data-component since novel insights are gained by monitoring and acting over a period of time. Where often such innovative machines or products are hardware-and software based, external companies enter the farm, supplying the hardware and the rudimentary software. In terms of the latter, one foreseeable business model is that of innovation by tapping into the software and the data generated by ‘smart’ hardware utilised on farms or greenhouses to improve already existing business models within a company. This can be done by the hardware providers themselves, or in combination with (groups of) companies such as a collective of farmers.

**Selling or otherwise opening up internal data to external parties**

Innovative business models can also stem from third-party software driven companies who can access this data in different ways. A myriad of earlier-described models exist, from freemium to subscription-based, that could be applied here. Remaining within the aim to improve existing business models, data generated within the agricultural company or within the network of agricultural company and hardware provider that can help improve the business model is opened up to be processed by an external party. This external party will highly likely be coming from the horizontal ICT sector to offer for instance data analytics services that can help improve hardware performance, maintenance timing or other intelligence that can be derived from a combination of internal data (for instance time of sowing against type of machinery used).
(iii) Using a mix of internal data and external data to create new business models
A third strand would be to find disruptive business models by allowing in external data to be combined with internal data to find new insights. A much-referenced example here would be John Deer, a company that moved from being a hardware manufacturer to becoming a data collector and broker, monetising value insights based on data collected over the years. Another example could be the combination of weather data with that of crop growth, allowing for customised and granular advice on where and where to sow, for instance. The question here is if and how the agricultural sector is really gaining from this, and how to measure the added value (in new information and insights, in a new (social) network, in machine maintenance, in administration and monitoring etc).

(iv) Cues on the current situation
Based on a set of exploratory interviews, some reflection on the current status of DDI within the EU agri-food sector. In terms of what is being done at the moment, both experts state that there are many possibilities, yet that the main activity still lies in the US and with large ICT players. Moreover, whereas there is a lot of potential for added value creation by combining data and performing (complex) analyses, Poulakidas points out that: ”Advanced data services and applications are rare. Even the simple things have not been done. There is a lot of Scope for IoT (in situ sensors) both in horizontal and vertical applications”. Also, they point out that often providers of the earlier-mentioned hardware for agri often keep the data (of maintenance schedules, and performance f.i.) to themselves. Charvat also explains “[...] the participation of big ICT systems and companies is very low. So SME’s dominate the market of agri-food. There is also a lot of tensions between big companies and SME’s due to awareness of the potential of this market”. The opportunities both experts identify can be found in GPS and satellite data, gaining more insights in crops over time and geography, in adding to sustainability and reducing environmental impacts and in insights into how the company’s day-to-day practice can be optimised. Farmers, however, often see the latter, as highly sensitive data, thus not easily given away or opened up to third parties. Even if data gathering

---

39 We have conducted interviews with Karel Charvat of the Czech Center for Science and Society on the 5th of May, who is an expert in Big Data for agriculture (See https://www.rd-alliance.org/users/karel-charvat) and a 2nd interview with Athanasios Poulakidas, who is the Coordinator for DataBio, a data platform of agricultural, on the 16th of June 2017. Interview questions can be found in annex B
D2.1: Report on high level consultation

is done for the company’s own use, there is a reluctance to invest in all sorts of sensors without being clear on a return-on-investment (ROI) scheme. Another connected challenges that they both identified would be inter-operability and the wide variety of ICT standards and harmonisation across Europe, both technologically and on the level of regulation. Within the EU, for now SME’s are getting engaged in DDI for agro, yet larger players do want to jump in due its potential.

4.4.2 Self-driving media centres: the changing landscape of transportation

Another sector of interest for the BVD-PPP is the transportation sector. This sector consists of many sub-sectors that are not easily comparable: from planes, boats and trains to drones, eBikes and self-driving cars, there is a lot of movement within the transportation sector. If we would need to categorise this sector, one way of doing so is to look at what is being transported. This would result in transportation of people and transportations of goods (animals and other living things also being in this category). In both sectors, DDI is very present in the terms of disruptive business models – in if we take a broad interpretation, transportation includes foodstuffs, building parts and materials, tourists, commuters but potentially also data as a good that is in transit. Recent examples of disruptive business models in this sector are numerous and are often connected to the concept of a sharing economy: from leftovers in the kitchen becoming a source of homemade-food delivery\(^\text{40}\) to Amazon becoming a main player in the transportation and delivery of practically everything.

In many of such examples, the service remains similar (f.i. food delivery or a cab service\(^\text{41}\)), yet the infrastructure to achieve the same goal or to fulfil the same needs has changed radically, also bringing about a complete different, data-driven way of adding value. Where such B2C and small B2B commercial services represent the forefront of DDI and start-up culture, another side of the transport sector is perhaps far less known for DDI among the general public, yet potentially even more impactful. When it comes data driven innovation in for instance large container shipping or public transportation hubs in urban areas, or the long-tail of the car manufacturing-and maintenance industry, there are major benefits and gains to be made, yet not so many DDI is taking place in such areas, as of yet. In the first set of examples, it also becomes easier to see how, due to the rise of smartphones and nearly ubiquitous connection to Internet within the EU, many media platforms can

\(^{40}\) See f.i. https://www.foodora.com/

\(^{41}\) See for example blablacar (https://www.blablacar.com/) or uber (https://www.uber.com)
example their services into other, previously unconnected sectors and places. A self-driving car can become media platform and office-on-the-go; a bike can become a health monitoring and air quality measuring device – the horizontal and vertical mash-ups of business models are not too hard to predict. Yet, in the more traditional sectors, actors are also moving and exploring how they can engage with DDI – ports are starting to become smart, for example\(^{42}\), and TLR London is opening up its API\(^{43}\). If there would be a common denominator, one aspect, or aim of becoming smart is to gain a better understanding of logistics on how are where goods travel from factory to consumer, or how people move from a to b – an almost inherently data-driven challenge. If we take a look at different areas of transport (be they public, private or commercial), where and how do we see emerging business models?

(i) Data to improve internal business models
One way of adding value to the transportation sector is to make use of internal data that so far has been non-or under-utilised. Highly dependent on the context within a sub-sector of transportation, such data would be available. In package delivery for example, tracking and tracing is already part and parcel of most service providers, offering insights from factory to shipping to day and time of delivery. Where such services are offered for free, the added fee is often found in delivery costs. Internal data on for instance public transport, however slowly opening up to the customer, yet a highly personalised service connected to the possible insights drawn from this data are rare. In other transport sectors, we need to delve into how and to what extend internal data is being used to improve existing business models.

(ii) Selling or otherwise opening up internal data to external parties
Companies and organisations in transportation (sub) sector(s) can also choose to open their internally collected data to third parties. Especially in the area of b2c applications, this has proven to be a fruitful direction of development. In the case of the Dutch train scheduling, for example, the leading train company had opened up (or accidentally had not closed properly) the scheduling information API, leading to a third party developing a widely successful train planner app. The company later on re-appropriated the app – but at the time was in no position to develop such an innovation in-house. In a similar vein, yet larger of scale and scope, is TLR London\(^ {44}\). This public transport company wants to not ‘just’ open up their APIs, they also plan

\(^{42}\) [https://www.bigdata-alliance.org/innovation-programs/smart-ports/](https://www.bigdata-alliance.org/innovation-programs/smart-ports/)

\(^{43}\) [http://www.bdva.eu/?q=node/716](http://www.bdva.eu/?q=node/716)

\(^{44}\) [https://tfl.gov.uk/](https://tfl.gov.uk/)
to offer various levels of entry to their data – some of it in one on one closed contracts with third parties, others for free and open to a larger set of actors. Regulatory challenges of public and open data aside (see section 6), the aim is to develop in-house their data infrastructure to such an extent that it allows for TR to offer different data products and services, practically becoming a data broker, data processor and analyst and data service provider.

(iii) Using a mix of (parts of) internal data and external data to create new business models
A continuation of the previous example would be for TLR to develop novel transportation services based on a combination of their own, internal datasets and eternal bought or otherwise utilised datasets that together lead to new insights and added value that would previously be impossible to achieve. To a certain extend TLR and other (semi-) public services are now starting to explore such models, also becoming active on data marketplaces and inviting startups and organising hackathons, for example, where the aim is not only to let other sin on large public datasets, but also to find new, cross – (sub) sectorial combinations. Linking train times to bus times but also to peak times in car transport in a region or city, for example, could lead to novel services, where companies themselves develop a data science department or a third party would be involved.

(iv) Cues on the current situation
With only a couple of examples covering part of the transportation sector, we will set out to provide a more in-depth overview of different DDBMs that are emerging within the Big Data Ecosystem. Based on interviews and a recent EU workshop on the matter, some key aspects are worth noting. One is that unlike the more commercial B2C types of disruptive DDI business models mentioned above, in the public sector, two experts point out that when it concerns public (service) data in Transport, there are many issues. One of them is homogenising datasets, as almost every member state has been developing and using their own ways of working. Another challenge is one of regulatory nature: the GDPR has been stretching the notion of personal data, making for instance location data personal data and even potentially sensitive data: how to draw the line between Personal and Non Personal

45 Presentation TLR at workshop APIs and business model, see http://ec.europa.eu/information_society/newsroom/image/document/2017-25/ws_8_june_shah_tfl_EE60E0CA-9D15-C665-13E1AB3D7AD24408_45580.pdf
46 Interviews with Vivian Kiousi, and Giorgos Panagiotopoulos, respectively Senior Research and Innovation Manager, Intrasoft International SA on may 3rd, 2017
D2.1: Report on high level consultation

Data is considered to pose many problems. The interviewees also mention that the problem of vested interests in the transportation sector: “For example the airlines want customers to move through Airports fast and thus they release data to make this happen. Airport shops on the other hand want people to hang around the shops so they are not interested in sharing data to speed up transit times.” Such competing interests often hamper sharing or free flowing of data, be they economical and/or political. In Vivian’s opinion there is a need to collaborate and to view Europe instead of pushing sectorial or national agenda’s. To do so we need to create excitement about the potential with citizens and governments. Open Data can be a valuable instrument to do this. For them the main trends in transportation are personalisation and tailored services. The challenge there for business models is to find a proper ROI, as often for now, just as in the agricultural sector, the added economic value is not entirely clear yet.

**Combi-case: Amazon moving into whole foods**

With the purchase of a whole-foods chain by Amazon, a horizontal data-driven actor is branching out to other sectors. The business idea behind this branching out is to expand on delivery - from books to goods to food, and more importantly, the network and data of whole-food shoppers. Horizontal players such as Amazon try to become the combining element in a lot of previously unconnected market branches.

The added value, and potential key - winning element will lie in the being able to obtain and make use of data about the network of collected actors and branches in the entire value chain of the whole-foods supermarket, from crops and livestock to in-store data analysis on customers to finally the delivery of goods and the eating habits at the dinner table. Data plays a role on multiple sides here for the development of this business model – it is not only about the physical chain that wholefoods represents, it is also the potential social and digital network that can be built up, making the data-players the spider in the web of supermarket services.

4.4.3 Summary

Considering the above, one of the main takeaways concerning DDI business models are that the added value within DDI does not necessarily lie in collecting mass data per se, but in finding valuable patterns and new forms of customisation and personalisation. This means that technologically, other foci gain relevance, such as Artificial Intelligence and Machine Learning. With this expansion of Big Data - related
D2.1: Report on high level consultation

techniques, also the amount and type of roles and types of activities expand (data retention, data marketplaces, data brokerage etc). With these new activities, also potentially novel business models emerge. Data Driven business models for now often come from within the ICT sector (horizontal), but there is a shift towards a vertical move of data driven innovation, as well as combinations of the two. The EU lags behind in data marketplaces, yet there are EU-based initiatives starting to emerge that want to offer EU-specific and EU-compliant databases. Data Science and Data Analytics for now is a relatively small field of expertise; in order for vertical novel business models to emerge and thrive, data scientists needs to be educated that have domain-specific knowledge on top of data science knowledge.

Where the analysis above is based on recent literature, we aim to develop and utilise a specific methodology in this WP to research where and how business models emerge within the EU Big Data ecosystem.

4.5 How to research business models in the Big Data Ecosystem

The main task of WP 2 is to provide a high-level regularly updated overview of the EU Big Data Ecosystem. In terms of business models and new forms of finding added value, we will follow the approach as given in the proposal. The following steps will be taken in order to research how and where business models are developing within the Big Data Value Ecosystem. As stated in section 2, per year the focus of WP 2 will shift to a different sector within the EU to explore the role of DDI within that sector.

For the first two sectors, Agriculture and Transport, in close collaboration with partners from the consortium, we will base our research methodology on the DAMIAN approach as the starting point and extend it in accordance to the requirements and particularities of DDI. Having analysed multiple methodologies that aim to capture the map, capture and foresight value creation in big data ecosystems47, the BVDA proposes to draw on insights from those methodologies and go beyond them. The DAMIAN approach investigates two factors of assessment, being a) the availability of alternatives and b) the level of openness to other players. In general, the approach consists of 5 steps that allow for a broad, if not complete capturing of all relevant aspects of an ecosystem, being:

47 See DAMIAN whitepaper
D2.1: Report on high level consultation

- Identify the problem-solution-space and scope;
- Mapping stakeholders, flows and assets in the data-ecosystem;
- Assess assets and determine dominant control points;
- Discuss impacts of control points;
- Discuss possible interventions and scenarios.

The key point of the DAMIAN methodology is that it specifically aimed at involving stakeholders to collectively map where added value might be found, but also where political, economic, regulatory and social-and ethical bottlenecks can be found, to subsequently develop direction for solutions to those bottlenecks. After phases 1 and 2, in which via desk research and interviews or surveys the problem-solution space and the accompanied stakeholders, data flows and assets have been mapped., the next step is to involve stakeholders in the research process. Via the use of a set of cards that depict core activities within the data process, accompanied by a set of questions, a value canvas can be drawn up that helps in mapping and understanding where an actor is within the data value chain, which dependencies and risks are connected to that position, and where and how next steps could be taken on that canvas. For the sectors transport and agriculture, the two sectors identifies as our target sectors for the first stage of the project, we are currently situated in phases 1 and 2, in which we utilise the network of the BVD-PPP to involve stakeholders in finding key drivers and barriers of DDI per sector. We are involving stakeholders in three different ways that semi-chronologically follow up on each other:

- Consulting stakeholders. For the information collecting on drivers and barriers for DDI within a sector, we rely on expert-knowledge of such a sector. Making use of the BVD-PPP, and in close collaboration with WP3, we are building a network of innovation intermediaries that we can draw from, and also can involve more actively in later stages in the project. Via interviews and report and paper by these stakeholders, we are in a constantly updating loop of what is going on ‘on the ground’. Making use of the lighthouse projects and iSpaces within the BVD-PPP, state-of the art innovation are caught, leading to a promised innovation radar that helps other stakeholders in the BDV-PPP network/ consortium.

- Involving stakeholders. Based on the analysis and first sketches of possible business case scenarios provided in section 4.4, more detailed and concrete application scenarios will be developed in close collaboration with experts. These application scenarios will be tested and validated during a series of workshops, of which the first are planned in November 2017. As a first of a series, such workshops are aimed to actively involve stakeholders in thinking about and developing sustainable business models for their sector or cross-
sector. Besides involvement via workshops (organised in close collaboration with WP3), we are inviting key stakeholders or actors within a sector or cross-sectoral to (co) author pieces for the BVD-PPP blog, ensuring accessible, high-level insights for other stakeholders.

- **Informing stakeholders.** The BVD portal and period reporting as well as a series of white paper will allow us to inform stakeholders, also linking them up to wider initiatives that might go beyond their scope or radar. Via special features (see section 2 on methodology) we aim to highlight a specific part of a section such as business models to offer a zoomed-in and in-depth analysis of a particular case. After having analyzed a set of cases, we aim to offer and inform stakeholders on different value creation mechanisms that can help develop the EU big data ecosystem.
5 Policy Special – Free flow of Data

5.1 Introduction

There is nothing inevitable about the impact of data-driven innovations. It will continue to impact society and economy, but how it does so depends largely on how European citizens, businesses and public authorities decide to use it and how the legislator shapes the regulatory frameworks for data-driven innovations. It is essential for the existing EU-businesses to grasp the opportunities of digital technology in order to remain competitive at a global level, and that for EU-start-ups to be able to scale up quickly with full use of cloud computing, big data, robotics and high speed broadband, thereby creating new jobs, increased productivity and sustainability. Being aware of these circumstances, on the 10th of January 2017, the European Commission adopted a Communication on “Building a European Data Economy”\(^\text{49}\) initiative. An integral part of the Digital Single Market strategy, the communication aims at fostering the best possible use of the potential of data to benefit European economy and society. The Digital Single Market strategy outlines the way for the EU to build a fair, open and secure digital environment.

The new framework covers the following objectives:

- improve access to anonymous machine-generated data;
- protect investment and assets;
- facilitate and incentivise the sharing of such data;
- avoid disclosure of confidential data, and
- minimise lock-in effects.\(^\text{50}\)

After adopting the Communication on “Building a European Data Economy”\(^\text{51}\), the European Commission published a working document on the free flow of data\(^\text{52}\). In this document the European Commission looked at the rules and regulations


impeding the free flow of data and present options to remove unjustified data location restrictions, such as legislative approaches to improve access to data. One of the legislative approaches is a proposal to *limit Member States’ ability to impose data localisation requirements*. In line with this proposal, the European Commission is currently preparing another proposal on *a new EU Free Flow of Data cooperation framework* as well as an *initiative on accessibility and reuse of public and publicly funded data*, as part of the Digital Single Market (DSM), i.e. Pilar I of the Europe 2020 Strategy. The new proposal is aimed at tackling “restrictions on the free movement of data for reasons other than the protection of personal data within the EU and remove unjustified restrictions on the location of data for storage or processing purposes”. It addresses the emerging issues of interoperability, usability and access to data in situations such as business-to-business, business to consumer, machine generated and machine-to-machine data. It would “encourage access to public data to help drive innovation”. In addition, the European Commission will continue its work on liability and other emerging data issues.

Restrictions such as geo-blocking are likely limiting the European Single Market for cloud-based applications and cloud services, among others. According to the estimates provided in “SMART 2013/0043 - Uptake of Cloud in Europe” report54, for the year 2020, cloud computing business could reach € 103b of net new GDP (including the public sector), a share of 0,71% of total EU GDP. Seamless flow of data in cloud services is seen as a key enabler of growth of the EU Data Economy in future years. Other important cloud aspects such as certification of cloud services and switching of cloud service providers are also closely linked to free flow of data and will be addressed by a specific European Cloud Initiative.

This section will analyse and discuss perspectives on free flow of data and its governance in the context of big data. It will illustrate potential impacts by reviewing applications in the Transport sector. Transport is one of the focus sectors in the first year of the BDVe project.
5.2 Enabling free flow of data

One of the key components is the Free Flow of Data (FFoD) initiative. Although, “the free flow of data in research across the EU, in particular access and reuse of research data that is generated by public budgets, is already having a positive effect scientific on collaboration and achievement of results”, the regulations are far from comprehensive and the free flow of data is severely restricted in many sectors.\(^{55}\)

Complementing the existing provisions in the General Data Protection Regulation\(^{56}\), the FFoD initiative aims to enable the free flow of data in the EU, through a legislative proposal removing unjustified data location restrictions. The FFoD initiative contains a mix of instruments and measures. Some initiatives such as Ending unjustified geo-blocking reflect the European Union’s strong free trading policies. In particular, the European Commission seeks to align the significant disparities in national (consumer) laws to which online businesses are exposed if they venture out of their domestic markets. The FFoD also includes measures to harmonise notice and takedown procedures for intermediary platforms, as well as measures to iron out unnecessary costs and processes involved in VAT and postal systems for online trade situations such as business-to-business, business to consumer, machine generated and machine-to-machine data.

One noticeable example where a comprehensive approach is lacking is the e-Commerce Directive.\(^{57}\) The Directive lays down the country of origin principle for information society services, but Member States are still permitted to impose restrictions on where data must be stored that limit the cross-border flow of data originating from that Member States. For example, there are restrictions from Member States on the location of data relating to financial and health sector, as well as company records, accounting and tax data, telecommunications and government data. The justification for such restrictions varies, with countries citing security, safety and other legal reasons.

Informed by a public consultation on the regulatory environment for data and cloud computing early 2016, the European Commission prepared an Ex-ante assessment of

---


\(^{57}\) Available at: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32000L0031:en:HTML.
D2.1: Report on high level consultation

the Free Flow of data initiative. The consultation highlighted the need to address data location restrictions, legal and technical restrictions to the free flow of data. It also highlighted the need to create new frameworks for data liability and to provide guidance on emerging issues around data access and associated mechanisms, data usage, and data transfers. Aside from the legal uncertainty of emerging concepts, such as restrictions on where data should be stored, are currently often constructed by contractual means; the report details four main problem drivers for the free movement of data:

- **Problem 1**: diverging data location restrictions and approaches in the Member States;
- **Problem 2**: unjustified or disproportionate data location restrictions in specific sectors or situations;
- **Problem 3**: the lack of European defined standards and practices on network, information security, prevention and investigation is causing data location restrictions;
- **Problem 4**: commercial users apply self-imposed data location restrictions in light of legal uncertainty and the lack of transparent requirements.

It is clear that data location restrictions are considered an important impediment to the free flow of data and consequently the establishment of companies wishing to deal with cross-border data transfers. Due to the legislative gap at EU level, Member States have applied their own rules and approaches that significantly diverge among the EU member states, leading to legal uncertainty.

Moreover, the report indicates that the focus of the regulatory intervention is directed at tackling data location restrictions and that – when mapping the options for the European Commission to take action – data location restrictions, fragmentation and weakness in sectorial policies would remain in existence without EU action. In the hope to solve this legal gap, the European Commission has proposed new regulations to remove unjustified data location restrictions with proposals on a wide range of issues that include:

- Geo-blocking
- Copyright reform
- Consumer law harmonisation for online commerce
- Intermediary liability
- European works quotas for online services

---

D2.1: Report on high level consultation

- Platform regulation
- VAT reform

Some of the key proposals are discussed below.

(i) Ending unjustified geo-blocking

On 25 May 2016, the European Commission tabled a legislative proposal to address unjustified geo-blocking and other forms of discrimination on the grounds of nationality, residence or establishment. The European Commission proposed a draft Regulation to ensure that consumers seeking to buy products and services in another Member State, be it online or in person, are not discriminated against in terms of access to prices, sales or payment conditions, unless this is objectively justified for reasons such as VAT or certain public interest legal provisions. The general objective of this proposal is to give customers better access to goods and services in the Single Market by preventing indirect discrimination by traders artificially segmenting the market based on customer’s residence.59

Next steps: the regulation is currently going through the European legislative process. Subject to approval by the Council and European Parliament, the Regulation is expected to be adopted in 2017.

(ii) A modern copyright framework

On 9 December 2015, the European Commission proposed a Regulation on the cross-border portability of online content services to allow EU residents to travel with content they have paid for in their home country.60 On 14 September 2016, the European Commission then published its proposed EU copyright reform package, comprising a proposed Directive on copyright in the Digital Single Market and a Regulation to adapt EU copyright rules to the realities of the digital single market. The proposals are intended to allow:

- better choice and access to content online and across borders
- a fairer and sustainable marketplace for creators, the creative industries and the press.

Next steps: the proposal is currently going through the European legislative process and must be approved by both the Parliament and the Council. There is no legislative timetable for this. It is also reported that the Commission intends to propose actions

59 Available at: http://ec.europa.eu/transparency/regdoc/rep/1/2016/EN/1-2016-289-EN-F1-1.PDF.
60 Available at: http://ec.europa.eu/transparency/regdoc/rep/1/2015/EN/1-2015-627-EN-F1-1.PDF.
D2.1: Report on high level consultation

to improve enforcement of all types of intellectual property rights, including copyright.

(iii) An analysis of the role of online platforms

On 25 May 2016, the European Commission published its communication on online platforms which:

- outlined the key issues identified in its assessment of online platforms; and
- set out its approach to online platforms in the future.

The Communication did not set out any legislative proposals, for example, by attempting to create a new law for online platforms. Instead the Communication set out the following guiding principles:

- create a level-playing field for comparable digital services;
- ensure platforms act responsibly;
- ensure transparency and fairness for maintaining user trust and safeguarding innovation;
- safeguard a fair business environment;
- ensure open and non-discriminatory markets in a data-driven economy.  

Next steps: the European Commission has announced a number of plans for this initiative. For example, it will reassess the need for guidance on the liability of online platforms when putting in place voluntary measures to fight illegal content online. By Spring 2017, the European Commission will also assess if further action is needed to address fairness in business-to-business relations. Most proposals are in the preparation stage, while some are in the process of being implemented by Member States. In a recently concluded mid-term review on the implementation of the Digital Single Market, the commission takes stock of progress on FFoD. In the next paragraph, we will discuss the review and its relevance for the BDV ecosystem.

5.3 Free flow of data and competitiveness

To remain competitive at global level Digital Single Market strategy launched by the commission in May 2015 aims to support businesses in the transition to a digital economy. Two years after the publication of the DSM strategy, the European

Commission issued a “Mid-term review on the implementation of the Digital Single Market” summing up progress so far. The mid-term review identifies where more efforts are needed and where the changing digital landscape calls for new action at the European level. The mid-review also describes several specific actions, including a legislative proposal to limit Member States’ ability to impose data localisation requirements; an initiative on accessibility and reuse of public and publicly funded data; an initiative to address unfair contractual clauses and trading practices by online platforms and guidance for online platforms on the notification and removal of illegal content and liability rules.

The focus is now on obtaining political agreement with the European Parliament and the European Council on all proposals, but above all the focus lies on the updated EU telecoms rules which are intended to boost investments in high-speed and quality networks, and that are critical for the full deployment of the digital economy and society. The mid-term review emphasises the political responsibility for the European Parliament and Member States to finalise key legislation and complete the strategy by 2018. The next 12–18 months are where the real benefits of the Digital Single Market strategy will (or will not) be converted from proposal to reality. The pressure is on the presidency of the Council of the EU to ensure that the legislative system actually supports - rather than stymies - the achievement of the proposals. Some of the key areas in which progress needs to be made are set out below.

**Promoting online platforms as responsible players of a fair Internet ecosystem**

Online platforms are considered by the EU as the “key gatekeepers” to the Internet ecosystem. In 2016, the EU identified two specific issues that it viewed as necessary for cultivating the growth of online platforms: (i) safeguarding a fair and innovation-friendly business environment, and (ii) ensuring that illegal content online is promptly removed from platforms. Since the initial analysis, the EU has conducted a fact-finding exercise on platform-to-business trading practices with results indicating some of the following prevalent issues: discrimination between different suppliers and sellers; online platforms’ self-preferential treatment of their own products and services; and a general lack of transparency into platforms’ practices. Consequently, the EU has promised to address issues such as unfair contractual clauses and trading practices by talking to relevant stakeholders and providing tools such as accessible dispute-resolution mechanisms, and fair-practices criteria. Fighting illegal content requires a similar approach, with buy-in from all stakeholders, in order to be successful. The EU has already committed to reviewing the need for formal, EU-wide flagging and removal mechanisms (so-called “notice
and action”), as well as the need for guidance on liability rules. However, there is scope for improving the ongoing dialogue between platforms, and the EU has promised to focus on technical solutions for the removal of illegal content, without undermining fundamental rights such as freedom of speech.

**Developing the European data economy**

It is estimated that the value of the data economy in the EU, if backed by a suitable legislative framework, will increase to €739 billion by 2020, and the number of data professionals will increase to 10 million by 2020 (from over 6 million in 2016). No doubt, the General Data Protection Regulation will provide critical support for the data economy, but it is no silver bullet, particularly for non-personal data. The EU recognises that a data cooperation framework will help to strengthen the cross-border free flow of non-personal data; this framework will address Member State’s’ legitimate interests on secure storage, while ensuring availability of data across borders for regulatory purposes. Coordinated action will also be needed to cover data issues related to cloud contracts for business users, as well as general enforcement powers to ensure that the free movement of data is properly implemented. In light of this, the EU has stated that it will provide a legislative proposal on the EU free-flow-of-data cooperation framework, to cover the principle of free flow of data, the portability of non-personal data and the availability of certain data for regulatory control purposes. The Internet of Things will certainly contribute to the exponential growth of the data economy, but significant support will be required to ensure a successful rollout in the EU. As a starting point, the EU will need to untangle the web of liability involved with defective connected devices and vulnerable software. To that end, the EU has stated that it will conduct further analysis on how it can effectively support the Internet of Things, beginning with a review of liability for data-intensive products, as well as the IP licensing framework that will be involved with such products.

**Tackling cybersecurity challenges**

Globally, cyber-attacks are already a very tangible risk to businesses and with an estimated 6 billion household devices being connected to the Internet in the EU by 2020, cyber-attacks will soon become an equally substantial risk to EU citizens. The EU has already recognised the potential pitfalls in this area by adopting a EU Cybersecurity Strategy in 2013, and the first legislative act (the Directive on Security of Network and Information Systems) in July 2016. However, the cybersecurity landscape is changing, and the EU has acknowledged that it is necessary to renew its approach here by reviewing the EU Cybersecurity Strategy and the mandate of
the EU Agency for Network and Information Security. Alongside this, there is a need to retain and develop essential cybersecurity industrial capacities. The EU has committed to developing measures on cybersecurity standards, certification and labelling to make ICT-based systems more cyber secure, particularly where they interact with connected devices.

To continue closed coordination between European Commission’s regulatory and antitrust enforcement agenda, the European Commission published the E-commerce Sector Inquiry. This report aims to identify particular online practices as problematic. One of its finding mentions big-data concerns, although it does not discuss this area in detail since big data were not a focus of the inquiry. Nevertheless, the European Commission notes that digital content providers depend on e.g. geo-blocking and on availability of licenses from content copyright holders and points to certain licensing practices that may make it more difficult for new online business models and services to emerge. Interestingly, however, the E-commerce report does not raise any red or even yellow flags in relation to geo-blocking practices that are addressed in legislative proposals as part of the DSM program. Both reports have made it abundantly clear that the EU needs to improve its dialogue with a variety of stakeholders to ensure that the right policy actions are made across all digital sectors in a resource-efficient and coordinated way to manage the digital transformation process and provide the “excellent infrastructure” that the EU believes is needed to underpin the Digital Single Market. The workshops that have mentioned in Methodology are examples of such process.

Another takeaways is the expectation that the Directorate-General for Competition will continue to focus enforcement resources in support of the Commission’s broader policy goals and that the EU has some way to go in order to achieve its DSM strategy goals. Against this background, the next section will discuss more in depth the challenges in the implementation of the Free Flow of Data initiative.

### 5.4 Challenges implementing the Free Flow of Data initiative

The 2015 Digital Single Market strategy contains highly ambitious targets also with regard to the FFoD. In the mid-term review, the Commission acknowledged the

---

implementation needs to pick up speed. One of the main challenges is to meet the aggressive deadlines set in the DSM strategy. A regulation enshrining the free flow of data across borders was expected in January and now it will not be released until the fall of 2017. Another point of contention is the Europe’s competitive position relative to the U.S. and China. The question is whether this will improve before all the measures in the digital single market strategy are implemented across all Member States. Self-driving cars are roaming U.S. streets under test conditions and the U.S. government has clear policy recommendations. The EU has a less concrete and coordinated approach. Specific actions in the fields of robotics and artificial intelligence have not been announced, despite a call for action from the European Parliament in February 2017. In addition to the concerns over the timelines there are a number of substantive concerns:

- **Copyright reform.** As well as taking on harmonisation of copyright exceptions the reform also targets geographic restrictions. It proposes “portability”, which means the possibility of taking a domestic content subscription across borders temporarily - or more aggressively a right of cross-border sale - allowing a licensee of, say Dutch rights to resell the content in Germany. A last minute addition stating that the Satellite Directive copyright regime should apply to online content threatens backdoor pan-European copyright. The Directive provides that copyright requires clearance only in the Member State from which the broadcast emanates; thereafter no other state’s copyright applies. The Commission claims that national copyrights will be upheld, but both geo-blocking prohibitions and cross-border copyright proposals clearly weaken a licensor’s ability to grant territorially limited rights. So it is still not clear how Digital Single Market will achieve goals without undermining national copyright. Moreover, the Satellite Directive reform would, if implemented, extend the directive’s de facto pan-European copyright regime to online services.

- **Platform regulation.** Another proposal is a review of the operation of major digital actors online. This appears to be an attempt to push for regulation of major search engines, social media or sharing sites. The proposal is at this stage only for a review. Some practitioners have questioned how sensible regulation of the “too successful” is as a legislative proposal. The main point is that the

---


65 A mixed bag of exceptions held clear by some countries, but with limited commercial impact.


D2.1: Report on high level consultation

European Commission implicitly assumes that digital technologies and the digitisation of non-digital business models will be relieved by the elimination of supply and demand constraints that result from insufficient e-communication infrastructure, fragmented copyrights systems, heterogeneous privacy laws, and simplified rules for e-commerce. It is true that legislation in these policies, if appropriately designed, can set a better context for digital economic activity. However there are questions regarding the enforcement of such laws. Some have pleaded for more grip on big data by competition authorities.\(^6^8\)

Within this context, the impact of Free Flow of Data initiative on key players in de BDVe has become necessary. We will address this in next paragraph.

5.5 Impact of Free Flow of Data initiative

The main upcoming FFoD plans of direct interest to PPP, are\(^6^9\):

- A legislative proposal on the EU free flow of data cooperation framework by Autumn 2017.
- An initiative on accessibility and reuse of public and publicly funded data, in Spring 2018.
- Further harmonisation of spectrum management and the strengthening cooperation with partners in third countries to develop common, open ICT standards for the Internet of Things, 5G, cloud computing and big data. Review the 2013 EU Cybersecurity Strategy and develop measures on cyber security standards, certification and labelling.
- Call on Member States to implement the strategy on Digitising European Industry and taking stock of results achieved by early 2018.
- Adopt a Communication in 2017 addressing the need and scope for further measures in the area of digital health and care.
- Promote Online Platforms as responsible players of a fair internet ecosystem.

These plans may impact the PPP, for instance on matters of collaboration. The Free Flow of Data initiative makes it possible for large databases to be made available, allowing for new ways of e.g. educating people and novel forms of interaction between people across cultural borders and/or improve societal insight in politics and the economy. Due to policy implementation and legal clarity, collaboration can


\(^{69}\) Available at: http://ec.europa.eu/newsroom/document.cfm?doc_id=44537.
D2.1: Report on high level consultation

improve by providing access to various data sources and/or individuals could be empowered by their new role as co-creator or co-innovator as well as generator and provider of data.

The next paragraphs will illustrate some of the impacts of the FFoD with the help of a real-world example: music streaming service of a connected-car. At this moment, there are restrictions for online subscription services. With FFoD, the EU is laying out plans to enforce new laws that break down geographical barriers for online subscription services like music streaming service in a connected car. When the new regulations of FFoD such as ending unjustified geo-blocking come into effect, online services must provide a service to its customers regardless of their current location, as long as they stay within the EU. For example, a person from Finland could buy a music subscription in their resident country, and then go abroad to Spain for holiday. The provider of the music streaming service would then have to allow the customer to use their same subscription in Spain with access to the same content that they get in the Finland. The rules are targeted at online video and music services, many of which currently region-lock their services to the user’s home country. The new policy would allow customers to keep using their music subscription anywhere in the EU. It gives customers more flexible about when and where they can use music and video subscription services (relevant to note here is that it does not mean that connected cars have to offer the same library of content across all EU Member States, just that if a customer signs up in one Member States, that same library can be accessed from any other Member States).
6 Conclusions

The first edition of this report focuses on methodology and setting the scene. It proposes to document the development of the Big Data Ecosystem across four dimensions: Markets, Policy, Architecture, and Norms. In the first edition, the focus is on Markets with a review of DDI business models and Policy, with a first review of the impact of the Free Flow of Data Initiatives. Future editions will zoom in on the issue of a common Big Data Architectures in Europe and the public perception of advanced Big Data analytics and AI (Norms).

**Markets: Data Driven Innovation (DDI) business models**

Summarising the main points from a business-model perspective, we can see a wide variety of levels of engagement with DDI and many different areas for progress and potential within each level. DDI strategies differ greatly per sector, and hold different models for value creation from activating by-product data coming from internal company processes to mixing and matching cross-sectional data sources to explore new connections. The more conservative sectors mainly aim to use DDI to improve what is already there, where the ICT sector behaves more horizontally, moving a role of historically being ‘support’ to a more centre stage position in the business model. Despite claims made in literature and by experts, there still exists reluctance in many sectors to ‘do’ DDI. Some of the main reasons are that ROIs are unclear, especially in DDI that requires higher levels of expertise and more specific sensor technologies. Another issue is that of waiting games\(^70\): before ROIs and real value of DDI proves itself, actors wait for ‘the other’ to make a first move. Horizontal (ICT) sectors seem to have the upper hand and leading the way in DDI, also bringing in most ‘disruptive’ ideas and cross-sector innovations. However, well-known models on the disruptive side of the ICT sector such as *freemium* might not automatically translate into workable and value-generating models in other sectors. Moreover, regulation ‘fights back’ against conflation and convergence of markets, for example, in a recent EU ruling that declared a digital cab hauling to be still a cab company, and not only a digital tech start-up. Such regulation might be interpreted as having a deterring effect on innovation. On the other hand, it also provides opportunities for EU-specific services regarding data retention, data marketplaces, and other data-services that up until now are heavily US-based or inspired.

\(^70\) See Robinson et al. (2012).
Policy: Free Flow of Data
The special feature on Free Flow of Data Policy looks at the status and impact of the upcoming Free Flow of Data legislative proposals and the Public data sharing initiative. With these proposals and initiatives, the European Commission aims at fostering the best possible use of the potential of data to benefit European economy and society. Within this context, the proposals and initiatives aim at enabling free flow of data and removing unjustified restrictions on the free movement of data due to the legislative gap, which is a consequence of application of different rules and approaches among the EU Member States. In the hope to solve this legal gap, the European Commission has proposed new regulations to remove unjustified data location restrictions with proposals on a wide range of issues that include for example: (i) geo-blocking, (ii) copyright reform, (iii) consumer law harmonisation for online commerce, (iv) intermediary liability, (v) European works quotas for online services, (vi) platform regulation and (vii) VAT reform. Most proposals are in the preparation stage, while some are in the process of being implemented by Member States. In a recently concluded mid-term review on the implementation of the Digital Single Market\textsuperscript{71}, the Commission takes stock of progress on FFoD. The focus is now on obtaining political agreement with the European Parliament and the European Council on all proposals to finalise key legislation and complete the strategy by 2018. One of the challenges is to meet this deadline and implement the FFoD initiatives. The implementation process may be complex and possibly slow given the broad impacts on current right holders. The next 12-18 months are where the real benefits of the Digital Single Market strategy will (or will not) be converted from proposal to reality.

7 References

**Journal Papers:**


D2.1: Report on high level consultation


White papers

Reports


D2.1: Report on high level consultation


portals, websites and blogs


http://www.bvda.eu
## APPENDIX A: List of interviews

<table>
<thead>
<tr>
<th>List of interviews</th>
</tr>
</thead>
<tbody>
<tr>
<td>Karel Charvát</td>
</tr>
<tr>
<td>Vivian Kiousi</td>
</tr>
<tr>
<td>George Dimitrakopoulos</td>
</tr>
<tr>
<td><strong>Planned</strong></td>
</tr>
<tr>
<td>Roberto Zicari</td>
</tr>
<tr>
<td>Corien Prins</td>
</tr>
</tbody>
</table>
Introduction

- What are key Big Data trends and developments in EU and elsewhere affecting Agri-food?

General free flow of data

- What is your opinion on the free flow of data discussion and how will it affect the Agri-food sector?

Opportunities and challenges

- What emerging opportunities do you see for big data in Agri-food in Europe?
- What are the main challenges in seizing these opportunities?
- How can these challenges be overcome?
- Who are the key players?
- What does the EU community need from the EC concerning big data opportunities and Agri-food?

Vision over 3 or 4 years

- What are key developments in the next 5-10 years for big data in Agri-food?
- What is your personal vision, going forward?

Background information
D2.1: Report on high level consultation

- What are the recent publications about big and open data in Agri-food?
- What are the recent events concerning big and open data in Agri-food?
**APPENDIX C: Interview Questions Transport**

### Introduction
- What are key Big Data trends and developments in EU and elsewhere affecting transport?

### General free flow of data
- What is your opinion on the free flow of data discussion and how will it affect the transport sector?

### Opportunities and challenges
- What emerging opportunities do you see for big data in transport in Europe?
- What are the main challenges in seizing these opportunities?
- How can these challenges be overcome?
- Who are the key players?
- What does the EU community need from the EC concerning big data opportunities and transport?

### Vision over 3 or 4 years
- What are key developments in the next 5-10 years for big data in transport?
- What is your personal vision, going forward?

### Background information
D2.1: Report on high level consultation

| What are the recent publications about big and open data in transport? |
| What are the recent events concerning big and open data in transport? |