D2.5: Annual Report on Opportunities

Workpackage:	WP2 – IMPACT: Framing the European Data Economy and maximize Impact
Editor(s):	Sonja Zillner (Siemens AG)
Responsible Partner:	Siemens AG
Contributors:	Sonja Zillner (Siemens AG) Tjerk Timan (TNO) Bas Kotterink (TNO)
Internal Reviewer	*Nuria de Lama, ATOS* *Andreas Metzger, UDE*
Status-Version:	V1.0 /Final
Due to	31/03/2018
Submission Date:	03/04/2018
EC Distribution:	Public
Abstract:	To stimulate and foster the growth of the European data economy, the continuous identification of data-driven opportunities is a critical task. In this deliverable we introduce a methodological framework for the identification and assessment of data-driven business opportunities that relies on the meta-analysis of demand and supply trends. We describe how the framework was

This document is issued within the frame and for the purpose of the BDVE project. This project has received funding from the European Union's Horizon 2020 Programme (H2020-ICT-2016-2017) under Grant Agreement No. 732630

developed and tested. We also provide a first example of how the framework can be used for analysing start-ups or SMEs. The framework description is complemented with an outlook to our future work.

Contents

List of Figures

List of Tables

Definitions, Acronyms and Abbreviations

Acronym	Title
AI	Artificial Intelligence
B2B	Business-to-Business
B2C	Business-to-Consumer
BDV	Big Data Value
BDVe	Big Data Value ecosystem
BDVA	Big Data Value Association
CAGR	Compound Annual Growth Rate
CPS	Cyber Physical Systems
DemoX	**D**ata-driven **E**cosystem **Mo**deling Approach
GDP	General Compound Product
ICT	Information and Communication Technology
IDC	International Data Corporation
IoT	Internet-of-Things
MS	Member States
SME	Small and Medium Enterprises

Table 1: Definitions, Acronyms and Abbreviations

Executive Summary

This deliverable addresses the question how data-driven business opportunities can be screened in an efficient and effective manner. Studies indicate that data economy growth relies on the successful matching of supply and demand. In particular the shift from technology-push to demand-pull is seen as important leverage criteria for market growth.

To support the process of identifying and scoping data-driven business opportunities by reflecting the dynamics of supply and demand trends, we are introducing the DemoX approach, which provides guidance in the continuous analysis of influencing factors on the demand and supply side. It is based on a conceptual model in the form of an ontology describing the aspects of supply and demand of data-driven innovation in a broad manner as well as guiding questions to provide guidance in exploring those aspects. The objective of the DemoX approach is to guide the process of analysing demand and supply dynamics of data-driven innovations by investigating the co-evolution and interactions between the scope of the offering (supply) and the context of the market (demand).

We leverage and combine existing business modelling approaches whenever possible, so as not to reinvent the wheel and to benefit from valuable established approaches. For instance, for analysing the central aspect of the data value chain on the supply side, the DemoX approach relies on a simplified version of the DAMIAN methodology. The various state-of-the-art approaches are combined and complemented with a meta-analysis for demand and supply trends to guide the process of identifying data-driven business opportunities.

This deliverable is the first of a series of reports that will be updated once a year. In this version, we are documenting the overall methodological approach encompassing the conceptual model / ontology and guiding questions. Based on this foundation, we are now able to analyse data-driven business opportunities in an efficient and effective manner. We plan to kick-start an easy communication with start-ups and SMEs in order to foster industrial engagement and investments. To indicate the value and relevance of the DemoX model, this deliverable reports on a first use of DemoX conducted by a data-driven start-up. This demonstrates how the DemoX categories are shaping and influencing data-driven business opportunities in entrepreneurial settings.

By relying on DemoX, we plan in the next year(s) to a) have a method in place that we can share with members of the BDV ecosystem for exploring data-driven business opportunities, b) analyse to which extent this content can be used for industrial training and university lectures by developing and evaluating content and guiding questions c) analyse patterns of successful start-ups to learn from them as guidance for industrial investment decisions and d) engage with SMEs and start-ups to help them scope promising business opportunities.

1 Introduction

The aim of this deliverable is to describe in some detail the approach by which the BDVe project will deliver on its goal *to document emerging business opportunities in the European Big Data Landscape* and to promote their uptake in key stakeholder audiences, in particular start-ups and SMEs.

The first part of the document (see Section 2) consists of a review of common methods to map and assess business opportunities such as the popular Business Canvas and methods proposed by Osterwalder (Osterwalder & Pigneur 2010)(Osterwalder et al. 2014). Here we also dive into the specific requirements of data-driven business propositions and the early methods designed to address them.

In Section 3 we describe and discuss a novel methodology (DemoX) that addresses the drawbacks of earlier approaches while preserving the ease of use of tools such as the Business Canvas. An intuitive application in line with the common business modelling practice of start-ups and SMEs will facilitate uptake by stakeholder intermediaries such as Business Incubators, Accelerators and Digital Innovation Hubs. Section 3 will also describe the results of the fist rounds of testing. In Section 4 will unveil the design of the stakeholder outreach based also on initial interviews with Big Data start-ups and we will conclude in Section 5 with a conclusion and outlook to our future work.

1.1 Approach

The main objective of Task 2.2 Opportunity Assessment is the meta-analysis of demand and supply trends of data-driven innovation. Our central approach to address this objective is to take one of the most popular business modelling approach among start-up companies and SMEs and extend and amend it to cover the reality of dynamic, (big) data-driven value propositions. This is what we call the DemoX approach. Once tested and validated the new approach will be introduced through the BDVe network of intermediaries (which we are building in Task 2.3 Innovation Booster) and to integrated the DemoX approach with the BDVe marketplace (created also in Task 2.3 Innovation Booster) to promote its use within PPP projects. To engage the wider community of Data-driven companies, in particular start-ups and SMEs we will create a portal within the leading European (and Global) start-up platform F6S. This will open up channels with start-ups already on the platform (1M+) as well as create synergies with the large number of start-ups already active in European programmes through the EU portal that F6S manages.

For the second phase of the project the aims is to link the DemoX-F6S portal to the BDVe marketplace by enabling a single point of authentication.

1.2 Impact and Timeline

The main impact of the work described in this report is to provide a continuous assessment of emerging, data-driven business opportunities by design a tool that can be offered to the main EU and national programmes addressing start-ups and SMEs that would be interested to validate or further the data component of their value propositions. The data collected through DemoX applications in different contexts will provide the intelligence to start monitoring and mapping an evolving landscape of European Data Value Propositions, following accepted and effective business modelling classifications.

The work on this task (2.2) started as planned in M7. The review of existing methods and the first design of an efficient, engaging and scalable method is concluded with the submission of the present report (M14-15). Next we will work on the wider testing and the finalization of the engagement strategy, in alignment also with relevant tasks in WP3 Community. By the end of Year 2 (M24) the aim is to have up to 100 companies submitting applications. In year three activities will focus on scaling up, applications, analysing and visualising results as well as refining and integrating the tool for large scale applications. By the end of the project we aim to achieve a continuous flow of applications of 100 companies per month across key domains and sectors.

As part of Task 2.3 Innovation Booster and 2.4 Industrial Investment we will work on the quantification of the impact in terms of competitive and additional attracted investment. This will require data gathering across the network of intermediaries. More detail on the impact measurement will be presented in the deliverables of Task 2.4.

2 Big Data Business Opportunities

How can data-driven business opportunities be effectively identified? This is the central question we aim to address in this deliverable. To address this question, we require an understanding of business opportunities in general as well as to know the characteristics of data-driven innovation. Those insights will indicate important aspects that require consideration while screening data-driven business opportunities and will guide the conceptualization of our methodological approach.

In the following section, we will first explain the concept of business opportunities; highlight the characteristics of data-driven innovation, as well as discuss the challenges of screening data-driven business opportunities. In this context, we will already sketch how the DemoX approach can be used for identifying data-driven business opportunities. The section concludes with an overview of state-of-the-art approach that describes in detail how we adapted the DAMIAN approach for analysing the supply part of data-driven business opportunities.

2.1 What are business opportunities?

Business Opportunity is a broad concept

In general, the concept of business opportunity is a very broad concept which is used to describe the chance to address a particular market need through the creative combination of resources that allow delivering advanced value propositions. The concept is used to describe a range of phenomena that begin in a early and immature stage and develop over time to more specific offering (Ardichvili et al. 2003).

Successful business opportunities align demand and supply side

In this way, the definition of promising business opportunities relies on the *balancing* of – often mainly technical – capabilities on the *supply side*, with the user needs and interests as well market dynamics shaping the *demand side*. (Timmons & Spinelli 2007) define *a business opportunity as something that has the quality of being attractive, durable and timely and that is anchored in a product or service which creates or adds value to this buyer or end user*. This definition implies two important aspects:

First, not every good idea leads automatically to a good business opportunity. This aspect is also often discussed under the label “technology search for solution”. For instance, Osterwalder intensively argues in his book (Osterwalder et al. 2014) that the technology-focused approach often slows down or hinders innovations as user needs and related impact criteria are only reflected in the very late phase of development. In other words, if the supply side capabilities cannot be mapped to a need on the demand side, business opportunities are not likely to evolve and grow.

Second, the window of opportunity, i.e. the time-interval given to implement and realize the innovation, is open as well as remains open long enough. This means that

for the respective entrepreneurs there is enough time to adapt a solution for the envisioned market demands. Or vice versa, in case entrepreneurs will not manage to deliver fast enough, e.g. due to strong external competition, we do not see this as a business opportunity.

Successful entrepreneurs never loose track of the demand side

In addition, studies indicate that most successful entrepreneurs and investors have a strong focus on business opportunities. In particular they continuously observe the demand side very carefully in order to understand what customers and marketplaces want and never lose track of it (Timmons & Spinelli 2007). Knowledge about market needs and concrete value proposition for specific user groups is aligned with the potential strategic positioning of those offerings on the market. The knowledge reflecting the demand side is used to guide the scoping of offering by combining own innovative technology components with reusable and available assets from others in a way that fosters competiveness.

Business opportunities are made and not found

The development of business opportunities is described as continuous process that involves proactive efforts to explore all essential steps of a new business. This process starts with simple concepts that are transformed step by step into more and more elaborated business concepts while entrepreneurs are investigating and developing them (Ardichvili et al. 2003). Thus, any tools and methods that provide guidance for such an ongoing process are seen as of high relevance.

2.2 Characteristics of Data-driven innovation

Increasing impact of data-driven innovation

According to (OECD 2015), data-driven innovation refers to the use of data and analytics to improve and foster new products and processes, new organizational processes and new markets and business models. In this context, data is seen as important resource, compared to water or energy that without preprocessing does nothing on his own. However, assuming that the right algorithms and technologies, data insights are in place, the usage of data is leading to better business decision, automation of manual tasks or automated orchestration of contribution, activities and tasks. Therefore, the role of *data and advanced data analytics* is getting more and more important in establishing the foundation for economic growth, development and well-being.

Data is an infrastructural resource

The *economics of data* has strong impact on the development of data-driven business opportunities. For instance, data can be consumed in unlimited number of times without losing its value, data can be reused as input for the production of different goods and services and -- on the other hand -- the value of data still depends on complementary assets related to the capability of extracting information

out of the data. Those fundamental economic properties of data, recommend treating *data as an infrastructural resource* (OECD 2015).

Data Economics drive disruption

Given the mentioned economic properties, *disruptions* through data are becoming more likely. In general, two different patterns of disruption can be observed: First, data-driven products and services transform the equation of offering and value / price and second, data-driven innovations harness network effects on various levels, such as access to data or customer as well as the **continuous matching of demand and supply** (Hagel et al. 2015).

2.3 How to screen data-driven business opportunities

Data economy is a high dynamic market

There are several factors that contribute to the *high dynamic of data economy*. *First*, the *European data market* which is measured by the value of the data products and services bought by European businesses and consumers *is rapidly growing*. According to the International Data Corporation (IDC)[1], the compound annual growth rate (CAGR) of the EU data market over the period 2016–2020 may be as high as 15.7% under the most favourable scenario. Those estimates imply that the size of the data market in Europe is expected to more than double in the coming years. *Secondly, recent technical breakthroughs* indicate a wide range of promising future business opportunities. For instance, Jeopardy demo by IBM Watson that could beat the top-two champions could demonstrate the high advances of expert systems, applications in speech recognition like Google Now or Siri are redefining human interactions, and image recognition and interpretation rivals with human performance (Davis et al. 2014). Experts estimate that even with no more technical advances in fundamental Data and Artificial Intelligence (AI) technology right now, it would take ten more years until we began to exhaust the commercial opportunities based on what we have right now[2].

Same, same but different

In addition, it is expected that the development of data-driven offerings will speed up as the existing data technologies along the data value chain are more and more reused, combined and aligned with each other. For instance, systems such as Watson that required development over several years with the involvement of a large team will in the future become available to ordinary software engineers. This in consequence leads to situations where entrepreneurs aiming to bring new

[1] IDC et al., European Data Market, SMART 2013/0063, D9 – Final Report, 1 February 2017, http://datalandscape.eu/study-reports

[2] Quote in Andrew W. Moore's presentation Developing AI strategy in 2017: what should universities do and what should industry do? February 2017

offering to the market require continuous scanning of the market's offering in order to identify promising available technology components, such as specific algorithms, knowledge models or hardware assets, that can be reused to speed up development time of their innovation. At the same time, entrepreneurs need to continuously investigate their own unique selling point and competiveness of their offering in a very dynamic environment. In such dynamic settings, innovations are no longer implemented by one organisation alone but rather a population of organisations and entrepreneurs that copy from each other as much as possible to ensure that technological assets can be reused and combined but that still put enough effort to ensure that they make a difference in the market with a unique offering. This can be compared to a swarm of birds flying into the same direction with each bird continuously observing where the others are flying to have enough distance to avoid collision but at the same time to be close enough to benefit from the wind shadow (Baecker 2007). In this way entrepreneurs need to continuously reassess what is part of their core offering and in which areas they are partnering with others in order to stay competitive in a fast moving market.

Matching of supply and demand is key success criteria for data market growth

In the comprehensive European Data market study (IDC & OpenEvidece 2017) three potential scenarios of how the European data market and economy is likely to evolve based on alternative assumptions regarding macroeconomic and framework conditions, indicated that the high growth scenario with estimated 4% of GDP growth between 2016 and 2020 is based on *supply-demand dynamics that shifts from technology-push to demand pull.* This can become possible through a fully developed ecosystem that is generating positive feedback-loops between data / technology companies and users. If there is a critical mass of technology companies that are able to develop innovative offerings that respond to high market needs and demands, the uptake (of technology provider) on the supply side is able to generate momentum, as -- due to the typical network effect of ICTs -- this rapid diffusion helps to multiply benefits for users /demand side.

Match-making of supply and demand

Thus, means that support the match-making between markets needs on the demand side with technical capabilities on the supply side help to stimulate the growth of the European data market. This match-making is not limited to classical end-users, such as for instance farmers who want to use predictive maintenance applications to optimize their farming task, but also happens between entrepreneurs or companies along the data value chain that would like to integrate available generic data analytics solutions, such as innovative chat-bot services or 3D rendering services, into a comprehensive offering addressing the needs of a particular customer segment.

In consequence, data-driven business opportunities that are described with a clear scope of offering per market segment (*supply side*) as well as reflect the ecosystem dynamics and benefits of network effects (*demand side*) are more likely to find a

promising market fit. Given the dynamics of the growing data economy, the relation between the scope of offering on the one side and the type of attributed value (e.g. price) on the demand side requires continuous reassessment. In consequence this leads to a *co-evolution between the supply-side (e.g. the offering) and demand-side (e.g. adjacent ecosystems)* for each data-driven business opportunities.

The DemoX Approach

The basic premise of the DemoX approach is that data-driven technologies are rapidly transforming not only the way how we live and work, but also how companies are organized and innovate, and the fundamental nature of a wide range of industries. This opens new perspectives, business opportunities but also challenges for the economic development. Therefore entrepreneurs, managers, policy makers and technology experts need to understand the changing technology paradigms, customer needs, business and ecosystem implications, and management practices of innovations in the data economy.

To address those needs we have developed the DemoX approach which is based on a **conceptual model in the form of an ontology describing the aspects of supply and demand in data-driven ecosystem.** For defining the categories we have built upon existing proven methods and theoretical concepts, such as (Osterwalder & Pigneur 2010), (Osterwalder et al. 2014) or *(Nooren et al. 2014),* as well as integrated well-proven established classifications, such as the BDV SRIA technical priorities (Zillner et al. 2017).

Our aim is to use the DemoX approach for two different purposes:

- First, we *want to engage with anybody* who is interested in practical insights of how to implement data-driven innovation in order *to guide investment decisions (time, money, etc.).* These can be students interested in founding their own start-up in this area, representatives of national contact points who can benefit from background knowledge to engage with industry, or decision makers from start-ups, SMEs or large corporates searching for a promising business opportunity. For this we have developed *a training concept with input and guiding questions* that guide the exploration of data-driven business opportunities. In order to ensure that the model brings value when applied to practice, we *decided to evaluate* it in a setting with non experienced innovators: We used the DemoX model as underlying guideline in the lecture "Data-driven Business Opportunities"[3] to guide interdisciplinary students to develop their own data-driven innovation. The evaluation of this lecture is part of our future work

[3] http://www.inno.tu-berlin.de/menue/teaching/winter_term_20172018/data_driven_business_opportunities/

- Second, we are interested in *identifying best practices or patterns* of data-driven business opportunities that can guide product, market and / or investment decision. Therefore, we are conducting a quantitative study to analyze data-driven business opportunities along with the categories of our developed DemoX framework. We use a representative sample of 90 startups from Crunchbase[4] and F6S[5] which have been coded along the DemoX categories to identify interrelations and patterns between the various dimensions as well as evaluate the importance of each category. The analysis of the sample set of start-up data is part of our future work.

In sum, we plan to use the framework for engaging with SME and startups in a way that allows easy communication. For instance, based on webinars we can share with interested start-ups and SMEs typical patterns of successfully implemented business opportunities in the data economy as guidance their strategic planning and investment decisions.

2.4 Related Work

2.4.1 Overview

To not reinvent the wheel, we aimed to reuse and combine existing business modelling methodologies whenever possible as well as to complement them with a *meta-analysis of demand and supply side trends* in order to guide the process of identifying data-driven offerings.

In our state-of the art analysis we investigated to which extent existing frameworks, research results as well as methodologies can be used to describe the supply as well as the demand side of the data-driven business opportunities. The DemoX approach builds upon popular existing business modelling methodologies and related research, such as (Osterwalder & Pigneur 2010), (Nooren et al. 2014), (Gassmann et al. 2014), (Hartmann et al. 2014), (Attenberger 2016) and (Johnson et al. 2008). In order to cover all dimensions we decided to align and combine them into a consistent model.

We could reuse valuable content from (OECD 2015) to scope the actors in data ecosystems and learn about characteristics and nature of data driven innovation in general. From (Adner 2006) we could reuse findings about the handling of risks involved when working with partners to either develop innovations or when engaging with partners required to adopt the innovation. We relied in our work on findings about emerging disruptive business and market patterns (Hagel et al. 2015) as well as insights about the different strategic roles for the governance of ecosystems (Iansiti & Levien 2004). In addition, we reused important concepts and

[4] https://www.crunchbase.com/

[5] https://www.f6s.com/

findings from research about emerging platform businesses, such as (Parker et al. 2016) or (Choudary 2015).

The aspect of data and technologies along the *data value chain* is a central aspect of the supply side of data-driven business opportunities. For exploring the data value chain, we relied on a simplified version of the DAMIAN methodology that we developed and prototyped in particular for the scoping of data-driven scenarios. This approach could be complemented with our findings in (Cavanillas et al. 2016) and with methodologies for exploring the value proposition (Osterwalder et al. 2014) and co-innovation partners (Adner 2006).

2.4.2 Damian Methodology

The DAMIAN method considers business opportunities as solution scenarios with an end-to-end perspective on service delivery: From observation through to analytics through to usage (to observation again). An end-to-end perspective ensures that all relevant technologies, drivers and barriers are covered for all actors along the value chain; socio-technical factors, vested interests, etc.

The methodology for the opportunity survey as detailed in Task 2.2 of the BDVe DoA draws on the TNO DAMIAN value-network analysis model (Nooren et al. 2014). DAMIAN was originally developed as a workshop-based tool to map, visualise and analyse alternative and competing end-to-end (digital) services in the converging domain of AV Services and telecommunication. In new applications it was extended to cover aspects of Data-driven Innovations (DDI).[6] DAMIAN distinguishes five stages in the design and evaluation of alternative DDI value propositions: (1) Defining the scope, e.g. datafication of container shipping services, (2) Mapping stakeholders and (data-) assets, (3) Evaluating key data-assets, (4) Identifying challenges, and (5) designing action or intervention (see figure below).

[6] Berkers, F. et al (2015), Understanding data ecosystem dynamics to accelerate value creation and mitigate risks, DAMIAN 2.0, TNO Whitepaper.

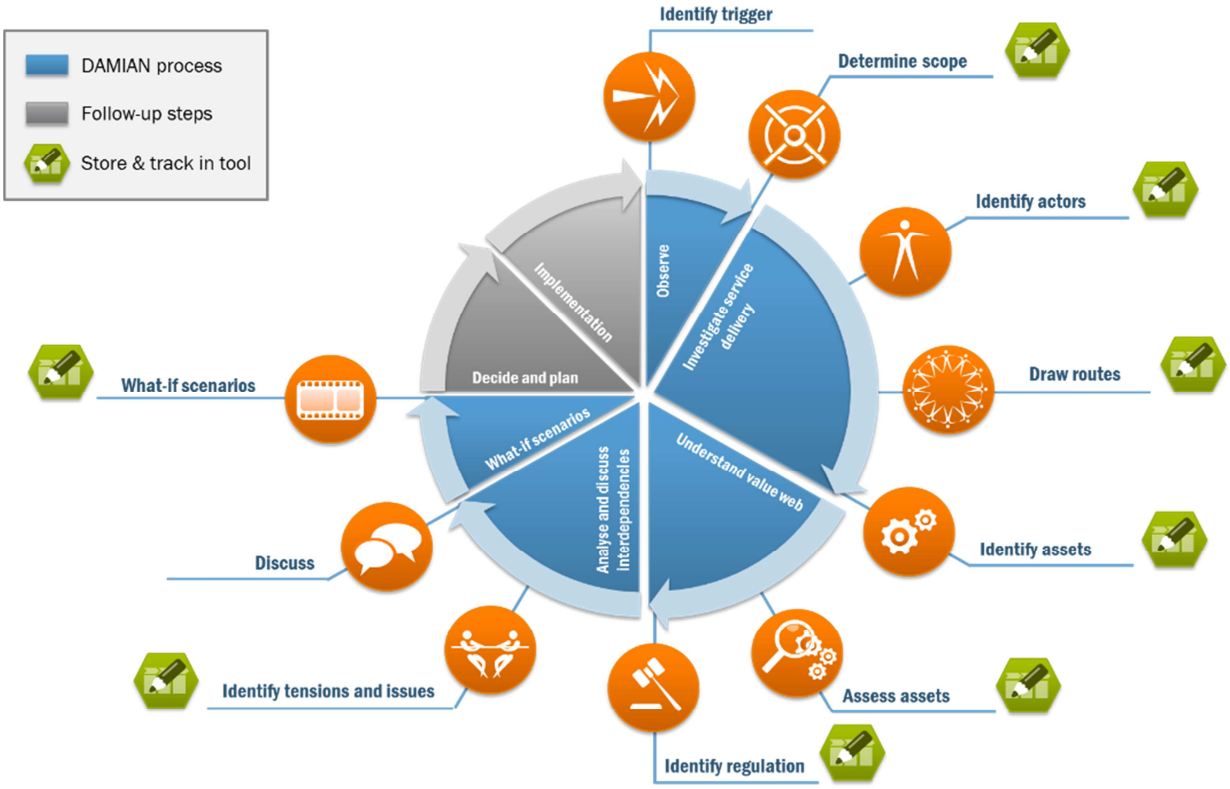

Figure 1 Overview of the DAMIAN process

Prototyping and testing the DAMIAN based approach on the identification of data-driven business opportunities revealed a number of drawbacks. Firstly, it appeared that any mapping of data-driven value propositions using DAMIAN required deep domain knowledge to be available to workshop participants. Another challenge for the participants was the level of complexity in the data collection cards. It was decided to further streamline the methodology to allow rapid implementation. For this the project team took cues from the success of the widely known Business Canvas Method (Osterwalder & Pigneur 2010). In the following sections we describe the results of the first pilot as well as suggested changes for an updated value network analysis approach.

2.4.3 Prototyping of Data-driven DAMIAN Methodology

Our goal was to develop a simplified DAMIAN Approach suitable to investigate data-driven offerings along the data value chain. The objective was to empower non-domain experts (the students) to explore data-driven innovation into an end-to-end manner as well as to empower them to validate it (i.e. identify challenges, solutions and compare differences (over groups) in different variants).

The starting point of the simplified DAMIAN methodology was a concrete scenario encompassing a persona with a concrete *value proposition*. In our case, this was the 35-year old Peter who requires Guidance and Navigation support for this scooter rides (see Figure 2)

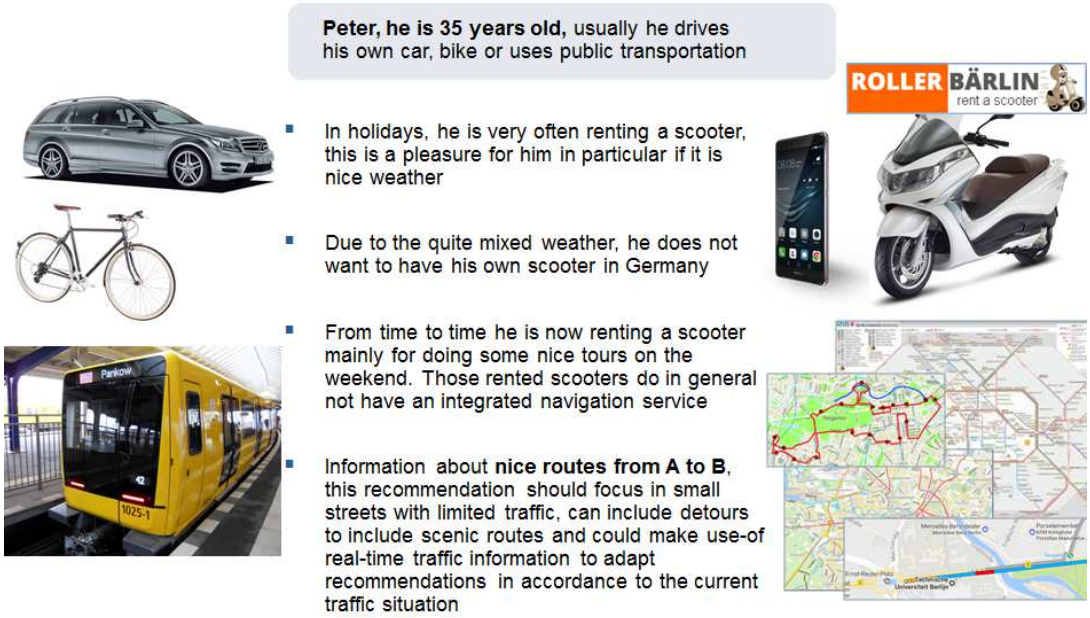

Figure 2 Description of persona and value proposition

In order to guide non-experts in the analysis steps, we relied on three steps:

1. Brainstorming to identify for a given value proposition all information about assets, such as data sources, technologies and partners and roles that will be required for implementing the data-driven innovation.
2. Mapping to display on the DAMIAN canvas in which steps and order the data and technologies, partner and roles are required for delivering the data-driven innovation
3. Assessment of the overall map of partners and technology to identify alternative paths, understand areas for outsourcing and ensure that the data-driven innovation is completely described in an end-to-end manner.

We prototyped a simplified version with approximate 60 graduate students from different disciplines attending the lecture Digital Innovation at TU Berlin[7]. In this prototypical evaluation, we learned that applying DAMIAN with an already concrete value proposition and persona in mind significantly reduces the complexity. Secondly, the students reported that the data-driven DAMIAN really helped them in scoping the very different assets and partners required for realizing the innovation in a systematic manner. However in the analysis it became clear that questions related to the roles and interest of players that could play an important role as cooperation partners were not covered within this DAMIAN version sufficiently well. In particular, the importance of investigating in more detail under which conditions partners would be willing to cooperate, was discussed.

[7] http://www.inno.tu-berlin.de/menue/teaching/archiv/summer_term_2017/digital_innovation

In sum we can conclude that the simplified DAMIAN methodology is a very promising starting point for investigating the supply side of data-driven innovations. We complemented the approach by a previous step to explore the value proposition along the approach from (Osterwalder et al. 2014) as well as a subsequent step to analyse co-innovation risks along the approach from (Adner 2006).

With the mentioned extension, we could establish a promising approach of analysing the supply side of data-driven innovations. However, the approach is still missing capabilities to analyse important dynamics on the market side, such as network effects, guidance for scoping value network strategies, and does not reflect the role of framework conditions.

3 The DemoX Approach

To reflect the economics of data, the disruptive potential of data-driven innovation as well as the high dynamics of data ecosystems, the screening and development of data-driven business opportunities requires a simpler framework that provides guidance in the *continuous analysis* of all *influencing factors* on *the supply as well as demand side*:

- On the *supply side* the focus is on the development of new offering. For a clearly defined value proposition, this includes the identification of and access to required *data* sources, the analysis of underlying *technologies* as well as functionalities and the identification of potential *co-innovation partners*, i.e. organizations or individuals that have access to complementary and needed capabilities to be in-sourced.
- On the *demand side* the focus is on the dynamics of the addressed markets and associated ecosystems. The analysis includes the development of a *value network strategy*, the analysis of partners in the ecosystem that are required to adapt the innovation in order to ensure market success (*adoption chain partner*) as well as the consideration of *framework conditions*.

To provide means for analysing the above aspects in an efficient way we have developed *the **D**ata-driven **E**cosystem **Mo**deling DemoX approach* which aims to support *the identification of data-driven business opportunities* in a continuous manner.

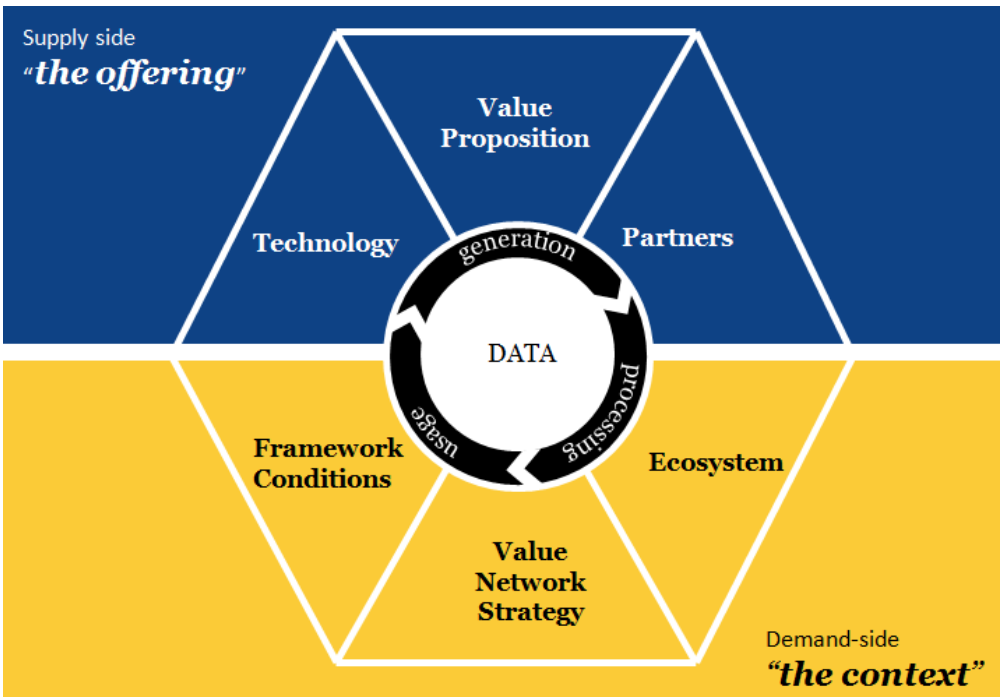

Figure 3 Overview of Data-driven Ecosystem MOdeling (DemoX)

3.1 Supply-side: Scope of your offering

The analysis of the scope of the offering starts with the scoping of your ***value proposition***, i.e. the definition of the business problem / user need you are planning to address. As the implementation of data-driven innovations relies on the use of data and data technologies, organizations need to involve third ***partners*** with their (often global) ***technologies*** to organize the value creation process along the data value chain. The various actors in the data ecosystem have different roles and interests that need to be reflected as basis for sustainable partnership building. DemoX guides in identifying the right data sources, selecting the necessary technologies along the value-chain as well as identifying co-innovation partners including risk assessments and strategies.

3.1.1 Supply Side: DemoX Ontology Dimension

3.1.1.1 Value Proposition

A value proposition describes why a customer should buy a product or service. In general it targets a well-defined customer segment. It allows solving a problem better than a competitive offering.

Due to the various data processing steps along the data value chain, including data generation, data storage, data aggregating, data enhancing, etc. steps, there are many different ways how players in the data ecosystem can produce value. In general we observe an increasing number of players, intermediary goods and services or technologies that are combined into comprehensive solutions for end-users. In the data ecosystem players are producing a potential intermediary value proposition that is again used by another company as input to develop another data-driven offering and so on. On the one hand, advanced technologies and practices are constantly emerging, and on the other hand, the existing technologies are combined and reused in a very dynamic manner. In order to classify the key offerings shared within the ecosystem, we rely and extend the classification of key actors in the data ecosystem in (OECD 2015). The *type of offering* describes the type value provided from an outside or demand perspective describing how an offering is positioned on the market.

This classification can be compared to the *scope of offerings* which is describing the components required to build an offering. With this supply perspective all additional assets for the key offering to work and for which potentially partnerships need to be established are analysed. For instance, TVision Insight[8] is a television measurement company that is pioneering how the true value of video content and advertisement can be determined and offers advertisement companies the analytical services measuring the attention level of the users. For being able to implement and realize such as service, they require to install a physical box which collects audience's movement data. Thus, the scope of offering includes a hardware asset that is never sold to anybody.

[8] https://www.tvisioninsights.com/

For describing the category "Scope of offering", we relied on the recent European data market study (IDC & OpenEvidece 2017) that is using four key components (Data, Software, Hardware and IT Services) to estimate the Data Market Value. We added a fifth category "solutions" in order to describe when different kinds of offerings are combined into a complex solution. Table 2 describes all important aspects to describe data-driven value proposition on a generic and abstract level as we could derive from (OECD 2015) and (IDC & OpenEvidece 2017)

Value Proposition		
Domain	**Category**	**Description**
Target Audience	B2B	In **Business to Business** (B2B) the addressed customer is a business customer. The Value has a stronger emphasis on financial values, such as reducing costs and/or growth. In case emotional drivers are of relevance, they follow economic value of offering.
	B2C	In **Business to Customer** (B2C) the addressed customer is an individual or household. The VP is typically a combination of functional and emotional benefits
Type of Offering	IT Infrastructure Offering	The market for **IT infrastructure offerings** encompasses both hardware and software offering. But most important for data-driven offerings are databases and related technologies and services (management, security, transport, storage) (OECD 2015). This also includes related hardware, for instance data generations, such as sensors, or hardware for automation of systems, such as drones or robots
	Data Offering	**Data offerings** enable access to data. This includes offerings from data brokers collecting and aggregating data, from public sector that provide access to data or even customers. Other examples are offerings from initiatives, such as MyData that are running infrastructures enabling controlled access to personal data. (OECD 2015)
	Data Analytics Service	**Data analytics services** are based on the foundation layer of IT infrastructures such as database, cloud and analytic software solutions. These offerings focus on specific analytical or visualisation solutions, targeting specific industries or even specialised tasks within an industry. (OECD 2015)

	Entrepreneurial solution	**Entrepreneurial solutions** address a concrete solution that can be applied by a concrete customer group. The main focus on user-driven solutions for end-users. (OECD 2015)
	Platform Business Offering	**Platform business offering** describes a plug and play mechanism that allows connected users and things to plug in and orchestrates them towards efficient interactions. It enables efficient social and business interactions that are mediated by software. (Choudary 2015)
Scope of Offering	Data	**Data** is acquired, acted upon, and sold within the data market. Corporate entities trade data internally in their organizations and between organizations.(IDC & OpenEvidece 2017)
	Software/ Analytics	**Software** associated with the data market is application software that falls under the business analytics product category or system management software. The business analytics tools included for analysis and delivery of data in the data market. In addition to application software is system management software — mostly used to manage hardware associated with the collection, storage, and analysis of data.(IDC & OpenEvidece 2017)
	Hardware	**Hardware** associated with the data market comprises server and storage for collection and storage of data, with some networking equipment included. Within servers most data work will involve high-end and midrange servers, so only a small share of volume server IT spend is included in this market.(IDC & OpenEvidece 2017)
	IT Service	IT spend in the data market associated with **IT services** varies over the development of the market. In the developmental phase, a large share of IT services spend will focus on training, education, and planning. In the later years this will transition to operations, maintenance, and support.(IDC & OpenEvidece 2017)
	Solution	**Solutions** represent combined offerings.

Table 2 Value Proposition

3.1.1.2 Data

The listing of data sources underlying a data-driven innovation guides us in understanding technical requirements and constraints, standard requirements, as

well as maturity of innovation. For that reason, any type of information that can be related to the underlying data source was documented in our study.
For identifying the data sources required for a data-driven innovation, the following questions can be reflected:

4. Scope / Subject: What data sources are required to implement the VP? Which real world objects are monitored? What information (e.g. data variables) are in this data? Why is this information relevant for the VP
5. Data Creation: Who is generating the data? E.g. a sensor, a process, an interaction, etc. Which actor is providing the sensors and/or software to monitor the physical world and virtualize the outcomes?

In our analysis we decided to rely on (for consistent analysis) extended the data types introduced by the Big Data Value Strategic Research and Innovation Agenda (Zillner et al., 2017). The analysis of data types guides us in understanding associated technical requirements, challenges and constraints of the required data processing technologies. Table 3 encompasses all important aspects to describe the various data sources. It relies on the list data sources described in (Zillner et al. 2017) which were extended and restructured to enable consistent representation.

Data		
Domain	**Category**	**Description**
Time and Space Dimension	Time Series	**Time series** is a series of data points indexed (or listed or graphed) in time order. Most commonly, a time series is a sequence taken at successive equally spaced points in time. Thus it is a sequence of discrete-time data. Examples of time series are heights of ocean tides, counts of sunspots, and the daily closing value of the Dow Jones Industrial Average.[9] Another prominent example of time series data is the behavioural data tracked with any user interaction on web or mobile devices.
	Temporal Data	***Temporal data*** can be time-series data or event data. In general, temporal data represents a state in time, such as the land-use patterns of Hong Kong in 1990, or total rainfall in Honolulu on July 1, 2009. Temporal data is collected to analyze weather patterns and other environmental variables, monitor traffic conditions, study demographic trends, and so on. This data comes from many sources ranging from manual data entry to data collected using observational sensors or generated from simulation models.[10]

[9] https://en.wikipedia.org/wiki/Time_series

[10] http://desktop.arcgis.com/en/arcmap/10.3/map/time/what-is-temporal-data.htm

	Geospatial data	***Geospatial data*** is data that has a geographic component to it. This means that the records in a dataset have location-related information tied to them such as geographic data in the form of coordinates, address, city, or ZIP code. GIS data is a form of geospatial data.
Different Data representation formats	Semi-structured data	**Semi-structured data** is a form of structured data that does not conform with the formal structure of data models associated with relational databases or other forms of data tables, but nonetheless contains tags or other markers to separate semantic elements and enforce hierarchies of records and fields within the data.
		A typical example of semi-structured data is **Web data**. Web data refers to the wealth of data stored in the www. Most prominent data source is wikipedia. When this data get extracted /crawled from the web, the data is in semi-structured format.
		Another typical example of semi-structured data is the wealth of **social media data** such facebook or twitter.
	Unstructured data	**Unstructured data** is free from data that is much more challenging to process as dedicated algorithm for extracting intermediate results from raw data are required. Each type of unstructured data relies on different and specific pre-processing algorithms and mechanism to extract metadata describing the content /semantics of the data. Often semantic models (meta-data and graph data are used to formally represent the extracted information from unstructured data).
		Text or language data and audio data refer to any type of unstructured text-based data, such as a radiology report, a case description or emails. Basically the text files encompass a long list of strings. Natural language techniques (NLP) are used to extract semantics from text documents. In accordance to the underlying language different NLP techniques have to be used.
		Image and Video data (Image data includes all mean for storing and transmitting photo graphic images as a digital file, e.g. tiff, jpeg, etc.
		Genomics data refers to the genome and DNA data of an organism. They are used in bioinformatics for collecting, storing and processing the genomes of living things. Genomic data generally require a large amount of storage and purpose-built software to analyse.[11]
		Other formats, such as laser

[11] https://www.techopedia.com/definition/31247/genomic-data

	Semantic Data	***Semantic data*** makes the content and semantic of data explicit and machine-processable. This includes ▪ **Graph data** refers to all formally represented knowledge. Most prominent examples are medical ontologies or the Google knowledge graph. ▪ **Meta-data** includes any type of data used for labelling or classifying other data sets. Metadata is defined as the data providing information about one or more aspects of the data; it is used to summarize basic information about data which can make tracking and working with specific data easier.
Combined /Complex Data	Internet-of-Things (IoT) Data	The **Internet of things (IoT)** is the network of physical devices, vehicles, home appliances, and other items embedded with electronics, software, sensors, actuators, and network connectivity which enable these objects to connect and exchange data.[12] IoT data refers to the exchanged data in IoT networks. When analysing IoT data several challenges need to be tackled: a) data structures as most sensors send out data with time stamps. Here the challenge is significant signals might happen only once in while but require then attention, e.g. through static alerts based on thresholds (Same challenges as time-series data) b) IoT data is often characterised by a strong correlation of time-series / sensor data and other sources of unstructured data, such as log files c) many machine learning algorithm require to aggregate sequential data into cross-sectional representation. This translation might cause lose of critical information.
	Media data	**Media data** encompasses all data sources used for media communications. This includes any type of communication channels through which news, entertainment, education, data or promotional messages are disseminated. This includes broadcasting and narrowcasting mediums such as newspapers and magazines TV, radio, billboards, direct mail, telephone, fax and internet. In this way, media covers a plural of mediums and thus very heterogeneous data sources. In the context of media data analysis there is a strong focus on video (e.g. youtube) and social media data (e.g. twitter, facebook), etc.

[12] https://en.wikipedia.org/wiki/Internet_of_things

Scope, Subject and origin of data	Personal data	**Personal data** is information relating to you only, which makes you identifiable – your name, photo, phone number, birth date, e mail address, car number plate, etc.
	Industrial data	**Industrial data** refers to any data assets produced and used in industrial settings of all areas. Often those are data produced from productions lines, energy systems, infrastructures, etc. In general, industry data is "closed data" meaning that it is "owned" (in terms of access rights) by the entity operating the product or machine or thing producing the data and that the data is likely to cover confidential information, e.g. a log file from a medical MR might contain valuable information indicating IPRs of the MR itself. This also includes **Operational data** in general produced in the context of operating any type of IT system
	Open data	**Open data** is the idea that some data should be freely available to everyone to use and republish as they wish, without restrictions from copyright, patents or other mechanisms of control. The goals of the open data movement are similar to those of other "open" movements such as open source, open hardware, open content, open government and open access.[13]
	Research data	**Research data** is defined as recorded factual material commonly retained by and accepted in the scientific community as necessary to validate research findings; although the majority of such data is created in digital format, all research data is included irrespective of the format in which it is created. [14]

Table 3 Data

3.1.1.3 Dimension of Data Value

Data value refers to the insights that can be generated out of data and how this can be used in a particular user or business context. In accordance to its value and complexity, we distinguish four different types of analytics (see Figure 4(

In addition, we observe that data is used to generate a lot of value for the automated orchestration of any type of consumers and producer, supply and demand, etc. Moreover, data analytics is increasingly used for automating existing processes, e.g. tasks originally accomplished by humans are replaced by smart algorithms

[13] https://en.wikipedia.org/wiki/Open_data

[14] https://www2.le.ac.uk/services/research-data/rdm/what-is-rdm/research-data

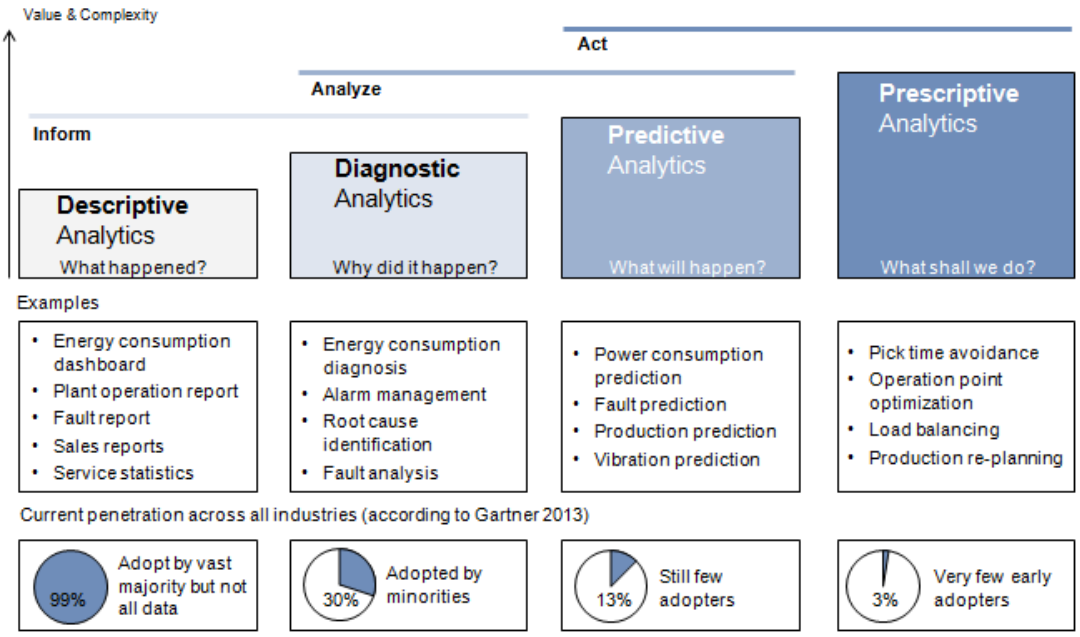

Figure 4 Value generated by the various data application

Table 4 encompassed the generic aspects of how data value is generated. The selection of categories is derived from expert interviews and an analysis of industrial presentations.

Data Value		
Domain	**Category**	**Description**
Insight Generation	Descriptive Analytics	**Descriptive Analytics** is the most frequently used analytics. Its main objective is to explain what had happened in the past by providing the analyst, business person or expert a view of key metrics that measure the area of interest. The traditional business intelligence and data mining applications fall into this category. They provide a very important basis for developing a deeper understanding of the underlying data sources.
	Diagnostic Analytics	**Diagnostic Analytics** aim to explain the root-case of a problem. Its main objective is to explain "why something happens". Those applications are often based on rule-based or semantic model capturing important background knowledge as well as flexible dashboards empowering the expert / user to explore or filter relevant features.

	Predictive Analytics	**Predictive Analytics** is about forecasting. Its main objective is to predict what will happen in the future, for instance the estimated point in time of a machine outage or forecasting a quantifiable amount of customer, etc. A predictive model relies on a variety of variable data that have a relationship which the event the model aims to predict.
	Prescriptive Analytics	**Prescriptive Analytics** is the basis for decision support and decision automation. Its main objective is to inform the machine or the user about the best course of action or strategy. Prescriptive analytics requires a deep understanding of the underlying engineering, business, mental or other processes in order to transfer analytics results into recommended actions.
Orchestration	Match-making	**Match-making** enables the automated orchestration of value generation. Match-making algorithms are mapping the demand side requirements with the supply side resource capabilities. Typical examples that make use of matching algorithm are dating platforms.
Automation	Process Automation	**Process automation** includes all applications that help to replace manual tasks or activities by machines or algorithms. This can range from applications that automate one very particular human tasks in comprehensive manner, such as supporting lawyers finding relevant cases, to applications enabling the automated orchestration of existing processes and workflow, such as AI-based planning and scheduling algorithms that develop strategies or actions sequences for execution by intelligent agents, autonomous robots and unmanned vehicles.

Table 4 Data Value

3.1.1.4 Technology / Data Value Chain

Similar to the use of the term value chain in the field of business management, the term data value chain makes the generic value-adding activities while developing a data-driven offering explicit. It highlights also that the data value chain consists of a series of subsystems each with inputs, transformation processes, and outputs. The data value chain can be used to model high-level activities that comprise a data-driven innovation. In (Curry 2016), the following four key high-level activities are mentioned: Data acquisition, data analysis, data curation, data storage, data usage.

For implementing the mentioned key activities, research in different technical areas is required. To ensure Europe's leading role in the data-driven world, the Big Data Value Association[15] has defined strategic and specific goals that are supported by key specific technical and non-technical priorities. In our analysis we will rely on the five technical priorities as defined in (Zillner et al. 2017) and represented in Table 5.

[15] www.bdva.eu

Technology / Data Value Chain		
Domain	**Category**	**Description**
BDV SRIA Technical Priorities	Data Management	The technical priority **Data Management** is motivated by the fact that more and more data is becoming available. This data explosion, often called "data tsunami" or "data lake", is triggered by the increasing amount of sensor data and social data, born in Cyber Physical Systems (CPS) and Internet of Things (IoT) applications. For more details we refer to (Zillner et al. 2017).
	Data Processing Architectures	The technical priority **Data Processing Architectures** is motivated by fast development and adoption of Internet of Things (IoT) technologies which is one of the key drivers of the Big Data phenomenon. Initially this phenomenon started by applying the existing architectures and technologies of Big Data that we categorize as data-at-rest, which is data stored in persistent storage. In the mean time the need for processing immense amounts of sensor data streams has increased. For more details we refer to (Zillner et al. 2017).
	Data Analytics	The technical priority **Data Analytics** aims to progress data analytics technologies for Big Data in order to develop capabilities to turn Big Data into value, but also to make those approaches accessible to the wider public. Data analytics will have a positive influence on all parts of the data value-chain to increase business opportunities through business intelligence and analytics while bringing benefits to both society and citizen. For more details we refer to (Zillner et al. 2017).
	Data Protection	The technical priority **Data Protection** addresses the need for advanced data protection and anonymization technologies in the areas of Big Data and data analytics. With more than 90% of today's data being produced in the last two years, a huge amount of person-specific and sensitive information coming from disparate data sources such as social networking sites, mobile phone applications, electronic medical record systems, etc., is being increasingly collected. For more details we refer to (Zillner et al. 2017).

	Data Visualisation and User Interaction	The technical priority **Data Visualisation and User Interaction** is addressing the need for advanced means for visualization and user interaction capable to handle the continuously increasing complexity and size of data to support the user in exploring and understanding effectively Big Data. Visual analytics is the science of analytical reasoning assisted by interactive user interfaces. Data generated from data analytics processes need to be presented to end users via (traditional or innovative) multi-device reports and dashboards which contain varying forms of media for the end-user, ranging from text, charts, to dynamic, 3D, and possibly augmented reality visualisations. For more details we refer to (Zillner et al. 2017).

Table 5 Technology / Data Value Chain

For the analysis of required (co-innovation) partner, we investigated in our qualitative study whether the technology required was available internally or provided by external partners.

3.1.2 Supply Side: Guiding Questions for Exploration

3.1.2.1 Value Proposition

- **Problem:** What is the business/user problem you are trying to solve?
- **User needs**: What is the sector and user segment you aim to address?
  - And what are the related user needs? (In B2C)
  - And what are the related (business) process needs? (In B2B)
- **Value:** What value are you bringing to the user?
- **Scope of offering**: What are potential offerings that will help your user get his / her job done more efficiently?

For exploring the subsequent aspects, we developed guiding questions complementing the data-driven DAMIAN Canvas:

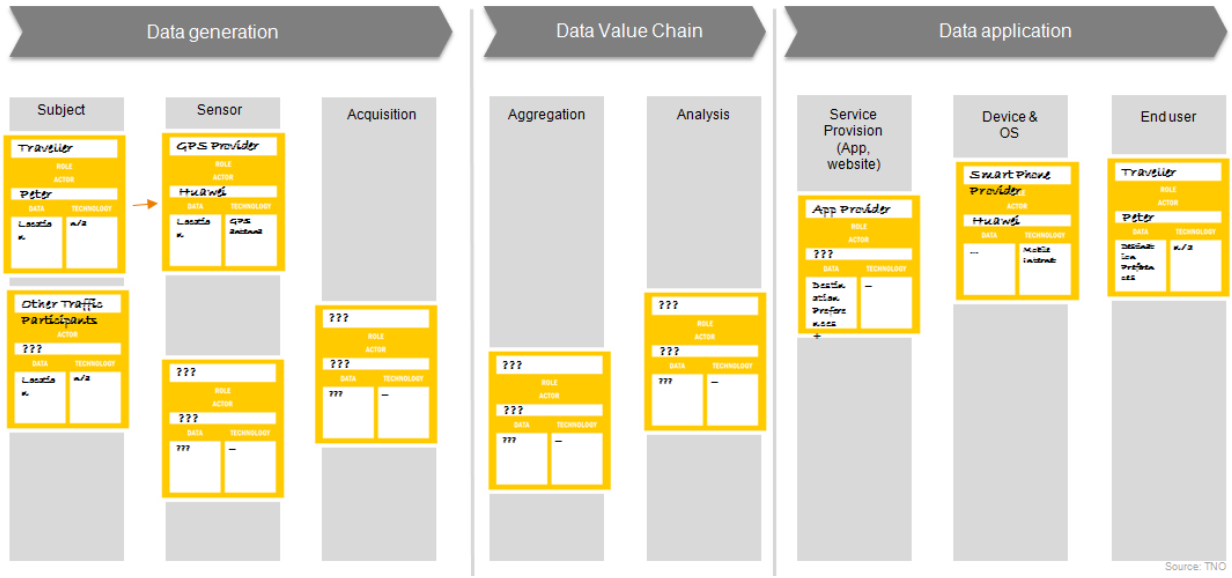

Figure 5 Data-driven DAMIAN Canvas

3.1.2.2 Data

Which data sources do I require? And to which data sources do I have access to?

- **Scope / Subject:** What data sources do you require to implement your VP? Which real world objects do you need to monitor? What information (e.g. data variables) do you expect in this data? Why is this information relevant for your VP?
- **Data Creation:** Who is generating the data? E.g. a sensor, a process, an interaction, etc. Which actor is providing the sensors and/or software to monitor the physical world and virtualize the outcomes?
- **Data Access / Acquisition:** How to access /acquire the required data sources? Which actor is transporting the data through its network? Which actor is acquiring the data sources through managed storage (e.g. cloud, local storage, dedicated server)

3.1.2.3 Data Value Chain

Which data processing technologies do I need for my innovation?

- **Data Aggregation**: How to combine and integrate multiple data sources? Which actor is aggregating multiple data sources through managed storage?
- **Data Storage / Management**: How to storage and manage the data sources? Which actor is managing the data in managed storage in accordance to the data processing requirements?

- **Data Analysis**: How to generate new data /information / new insights based on the available data? Which actor generates new data/information based on the collected data? Which software/hardware is used to analyse the data?
- **Data Visualisation**: How to visualize the data / information generated through the preceding steps? Which actor is providing the software (e.g. apps, website, algorithm) to explore and visualize the data / information?

3.1.2.4 Data Application

How to realize / implement the data-driven innovation in the real world?

- **Delivery:** How is the new offering delivered to the customer? E.g. app, website? Which actor is providing the portal or platform (e.g. app store) to access the data/information?
- **Device / Hardware & OS**: Do customers require particular hardware to access/use the data-driven offering? Which actor is providing the hardware, e.g. smart phone, tablet, PC, TV) to access data/information?
- **Real-World-User**: Which actor can make decisions based on the generated output? Which main tasks are supported by the output? Which main benefits are provided to the user?

3.1.2.5 Co-Innovation Risk Analysis

Step 1: **Identification** of Co-Innovation partners and their contribution

- What human, data source, data processing, technical, financial, etc. capabilities are required to create the envisioned data-driven innovation?
- What organizations or individuals have capabilities we do not have? What is our relation to them currently?

Step 2: Co-Innovation **Risk Assessment**

- How important is the contribution of the co-innovation partner for the success of our innovation?
- What are potential risks and effects that we need to keep in mind?

Step 3: Identify the various **strategic options** to address the co-innovation risks

- **Resources**, e.g. can we add resources (money, talent, or both) to bolster development effort?
- **Scope**, e.g. can we focus on a more modest offering?
- **Timing,** e.g. can we opt for a less aggressive timeline?
- **Partnering**, e.g. can we have several partners working on the same challenge?

3.2 Demand-Side: Context of your market

The demand-side analysis covers a sound understanding of your ***value network strategy***. Related questions focus on identifying stakeholder group(s) that will pay for an innovative solution, the combination of the various revenue streams and how they will change overtime. This is complemented with a long-term perspective of a promising strategic position in evolving ecosystems. Complementary, the aspect of the ***ecosystem*** covers adoption chain risks and strategies to overcome them (Adner 2006). Finally the aspects of framework conditions analysis supporting as well hindering ***framework conditions***, such as free flow of data, or (open) standards, will be reflected.

3.2.1 Demand Side: DemoX Ontology Dimension

3.2.1.1 Value Network Strategy

A value network strategy is a way to understand all of the market components and dynamics that come together and form the basis for positioning a data-driven innovation that adds value to an intermediary player or the end-user.

In the data-driven economy, value streams are no longer bi-directional but involve several players exchanging different types of value. The party who is benefiting from a value-added service, no longer needs to be the one who is paying for the service. Such value networks already exist in the internet environment.

For example, most of the established players providing data-based solutions, such as Google, eBay, YouTube, Facebook, ITunes, etc. are building up a growing user community by offering them free services, which allows them increase their income as each advertising companies is paying a fee per click or user.

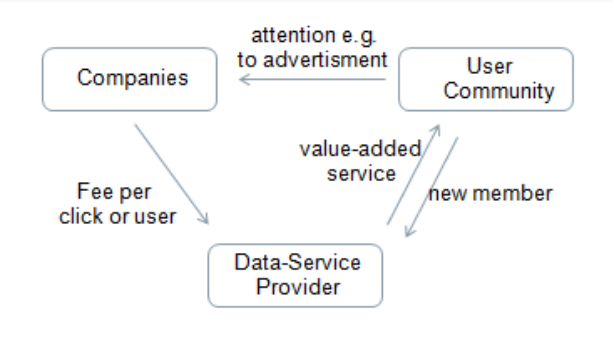

Figure 6 Value streams of Internet-based Services

This category reflects the impact or trajectories of the envisioned value proposition in the market(s). You can see it as the 'ecosystem version' of the value proposition that explains how the value proposition is positioned in the market. Here we investigate the dynamics of (an adjacent) platform business, explore how to make use of network effects and analyse whether the value proposition stacks up against those of competitors but also how it fits in with potential strategic partners in the

value network. In addition it addresses the question of positioning a particular value proposition given the reality of existing business models and strategies in the markets by scoping a sustainable capture value strategy.

The Value Network Strategy consists of three different aspects to be discovered in further details. First, the organisation needs to understand which *strategic role* it aims to play in the context of opportunity, as well as secondly, how to *harness* underlying *network effects*. Finally, for bringing the mentioned first three aspects into a consistent perspective, the organisation needs to develop a roadmap of revenue models explicating how the various *revenue streams* can be combined and adapted over time to ensure sustainable value capture.

3.2.1.1.1 Strategic Role in data-driven Markets

From our practical experience, we rate this as a very important aspect, because the way how companies are organized should rely on a clear strategy aligned with the product focus and market impact. To best of our knowledge – we assume that this is due to the fact that this topic is only emerging – there exist not agreed upon classification of ecosystem strategies. This might due to the fact that strategies can only evolve as soon as –often disruptive market patterns – have been discovered and described.

Due to its high relevance, we decided to cover this important aspect by analysing strategic roles;

- in already *established data-driven ecosystems*. Here we refer mainly to the comprehensive analysis of (Iansiti & Levien 2004) distinguishing keystone, dominant and niche player. The actors in such ecosystems usually differ in the degree they can influence the other members of the ecosystem.
- in *traditional markets*. Here the main focus is to transform traditional businesses and services into digital and data-driven services. This approach is also known under the label Smart Data[16].
- in *emerging and potential disruptive market set-ups.* Here we refer to relevant disruptive patterns described in (Hagel et al. 2015), such as the strategy of creating markets by connecting demand and supply in an innovative manner or the strategy to position offerings as platform that establish a foundation for others to build upon.

No governance strategy is better than any other. However, a company should be clear about how to set up his own cooperation strategy within the ecosystem. Table

[16]https://www.siemens.com/global/en/home/products/mobility/rail-solutions/services/digital-services/smart-data.html

6 summarizes the main strategic roles we could derive from the literature as well as analysis of start-ups.

Value Network Strategy: Strategic Role		
Domain	**Category**	**Description**
Established data-driven Ecosystems	Keystone Actor	**Keystone actors** shape and coordinate the ecosystem largely by dissemination of platforms that form a foundation of ecosystem innovation and operations. As keystone player profit from the number of members in the network, they usually take responsibility for their well-being and survival. (Iansiti & Levien 2004)
	Dominator Strategy	In contrast to this is the **dominator strategy** that attacks the ecosystem by absorbing and integrating external assets into internal operations. However, empirical findings show that companies who have dominant influence over the business ecosystems' key resources and at the same time do not take responsibility for the long-term health of the associated (often small) ecosystem partners easily jeopardize the economic basis of the whole ecosystem. (Iansiti & Levien 2004)
	Niche Strategy	A large number of organizations follow a **niche strategy**. Those companies emphasize differentiation by focusing on unique capabilities and leveraging key assets provided by others. (Iansiti & Levien 2004)
Traditional markets	Smart Data Business	**Smart data business** investigates the usage of data to improve traditional business outcomes. This ranges from industrial applications, data-driven marketing to process automation. Its main objective is to pre-process, analyse and contextualise data in a way that decision making and business functions can be improved.
Disruptive settings	Connect Supply and Demand	The objective of the **Connect supply and demand** strategy is to expand marketplaces by connecting fragmented by producers/sellers and consumers/buyers. With new distribution channels it becomes possible to attract more consumer with a wider range – often personalized – offerings than was previously possible. (Hagel et al. 2015)

Table 6 Value Network Strategy: Strategic Role

3.2.1.1.2 Strategies to harness Network Effects

Data-driven innovation very often benefit from network effects. In this category we aim to understand on which level benefits are emerging. A network effect occurs when a product or a service becomes more valuable to its users as more people use it. (Shapiro & Varian 1999). Table 7 describes the three different levels of network effects derived from (Choudary 2015).

Value Network Strategy: Network Effects		
Domain	**Category**	**Description**
Network effects	On data level	With higher amount as well as more different types of data sources, the envisioned value proposition can be improved. I.e. with more data, the underlying algorithm can be improved as they become more accurate. Typical Examples here fore are TomTom or Google applications. In addition, data is used to improve the matchmaking between demand and supply.(OECD 2015) (Choudary 2015)
	On infrastructure level	**On the infrastructure** level the availability of tools, services and rules that enable the seamless integration of technology components as well as automated orchestration of processes value creation. This plug- and play nature of the tools and services establish the basis for others, external companies and producers, to build value and services on top of this infrastructure. (Choudary 2015)
	On marketplace	Network effects **on marketplace level** describe a situation in which the increased and balanced number of producer and consumer on a marketplace increases the value of the overall marketplace. It is based on the participants on the platform and their relationships. (Choudary 2015)

Table 7 Value Network Strategy: Network effects

3.2.1.1.3 Data-driven Value Capture: Revenue Models

As of today, data-driven revenue models are rarely covered in related scientific research. This might be due to the fact that organizations in general do not very openly talk or write about their revenue models (Attenberger 2016)

When analysing successful data-driven innovations, we can observe that the single revenue models used have not changed significantly through the emergence of big data technology. However, the major and very important difference of revenue streams for data-driven innovation is that

- Data-driven innovation relies on very different revenue streams
- in general each customer /stakeholder segment is addressed with at least one revenue model
- those revenue streams are very likely to change over time (in accordance to market dynamics, competitive situation, the successful building up of a customer base (or provider base in case of multi-sided market, or the market adoption)
- the set of all revenue models are all conceptually and over time aligned with each other similar to a roadmap

For example, the revenue model approach of the US-based company of **Practice Fusion** offering **innovative EHR technology** illustrates this approach. Usually, enterprise software for medical practices can cost $50,000 and more. However, thousands of physicians receive free electronic health record from Practice Fusion. How can this be?

Practice Fusion bases their revenue model on various pillars which are combined as follows (Anderson 2009):

- **Freemium** revenue model: Practice Fusion offers a free software version including **advertising** that is aiming to attract critical mass of users, in this case healthcare provider. The same software is also offered without advertisement for a **subscription**-fee for 100$ per month for.
- **Selling access to data**: Medical associations conducting research on specific conditions require longitudinal health records for a large set of patients:
  - one patient's anonymized chart: $50 to $500
  - can be sold multiple times

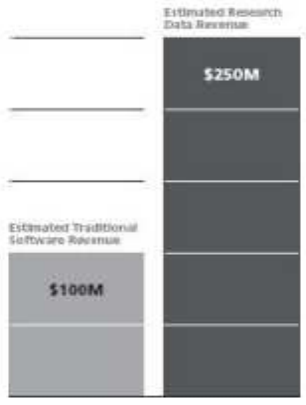

Figure 7 Revenue stream calculation for Practice Fusion EHR technology (Anderson 2009)

Table 8 consolidates and describes a list of data-driven revenue models that we derived from (OECD 2015), (Osterwalder & Pigneur 2010), (Hartmann et al. 2014).

Value Network Strategy: Revenue Model		
Domain	**Category**	**Description**
Data-driven Revenue Model	Freemium	The term **"Freemium"** is a combination of "free" and "premium". According to Hartmann et al., 2014), Freemium is the most dominant revenue model in the data ecosystem in particular for start-ups in the B2C area. The general idea is to provide data-driven offerings free of charge but charge money for additional features of the offering (premium). In addition, Freemium model is often combined with the advertising-based revenue model
	Advertisement	Here the revenue stream results from fees for **advertisement** of a particular product, service or brand. This means, people can consume data-driven offerings free of charge or with a discount in exchange for viewing paid-for advertisements (OECD, 2014d). Increasingly, advertisement is provided based on the profile and/or location of the consumers. Advertisement-based revenue models are also very common in multisided markets with a minimum of two distinct user groups that synergically support / complement each other.
	Subscription	Revenue is generated by selling continuous access to a service. Examples of **subscription-based models** include regular (daily, monthly or annual) payments for access to the Internet, as well as access to digital content including data, news, music, video streaming, etc. The category also includes regular payments for software services and maintenance, hosting and storage, and customer "help" services.
	Usage Fee	Revenue is generated by the use of a particular service; the more service is used, the more the customer pays. **Usage fees** are typically charged to customers for use of a particular (online) service – including most offers that are provided "as-a-Service" (XaaS), such as cloud computing based services. These services are offered through a pay-as-you go model, where usage fees are charged for the actual use of the service.

	Asset Sale	Selling of goods (including digital content) – **Asset sale** is still used in the data ecosystem, mainly by IT infrastructure providers. But it is also used by service platform providers that sell sensor-equipped smart devices (including smartphones, smart meters and smart cars) as a source for generating data and delivering value-added services. Furthermore, it includes pay-per-download revenue models where users pay per item of download. These could include, for instance, data sets or other digital content such as e-books, videos, apps, games and music.
	Licensing	This revenue model is often used to generate revenues from intangible assets that are protected through intellectual property rights (IPRs), such as patents and copyrights. Licensing may thus be used to monetise software and software components including algorithms, libraries and APIs. It may also be used for databases.
	Selling of Services	This revenue model includes the provision of traditional **B2B services** such as IT consultancy services, software development and maintenance and helpdesk support. It also includes a wide range of long-term B2B services provided by Internet intermediaries such as web hosting, domain registration, and payment processing. It thus overlaps with the revenue models that are based on subscriptions and usage fees often used for IT service contracts.
	Commission Fee	This is mainly used in B2C markets by intermediaries that use data analytics to better match supply and demand. Payment often will be calculated on the basis of a percentage of the price of products supplied, and it will only be obtained when successfully matching supply and demand – that is, when successfully providing businesses with customers.

Table 8 Value Network Strategy: Revenue Model

3.2.1.2 Framework conditions

Framework conditions identify the main factors which will enable or prevent the development of the European data market and economy. They can be divided into policy /regulatory framework conditions on the one side and market /non-regulatory conditions on the other side (IDC & OpenEvidece 2017) (see Table 9).

Framework Conditions		
Domain	**Category**	**Description**
Policy and Regulatory	Copyright	Digital technologies have radically changed the way creative content is produced, distributed and accessed. European Commission is adapting the EU copyright rules to new consumer behaviours in a Europe which values its cultural diversity. **Copyright** ensures that authors, composers, artists, film makers and other creators receive recognition, payment and protection for their works. It rewards creativity and stimulates investment in the creative sector. 33 sectors of the EU economy are considered copyright-intensive, accounting directly for over 7 million jobs, or 3% of employment in the EU[17].
	Data protection and Regulation	The protection of your privacy, including your personal data, is of great importance to the European External Action Service (EEAS)[18]. The objective of this new set of rules is to give citizens back control over of their personal data, and to simplify the regulatory environment for business. The **data protection** reform is a key enabler of the Digital Single Market which the Commission has prioritised. The reform will allow European citizens and businesses to fully benefit from the digital economy[19].
	Free flow of data	**Free flow** of non-personal **data** is a pre-requisite for a competitive data economy within the Digital Single Market. To fully unleash the data economy benefits European Commission aims have policies in place to ensure a free flow of data, allowing companies and public administrations to store and process non-personal data wherever they choose in the EU[20]. This aspect is related to the discussion and action regarding the ownership of data as well as control of data access and usage.

[17] https://ec.europa.eu/digital-single-market/en/policies/copyright

[18] https://eeas.europa.eu/headquarters/headQuarters-homepage/3032/transparency-and-data-protection_en

[19] http://ec.europa.eu/justice/data-protection/

[20] https://ec.europa.eu/digital-single-market/en/free-flow-non-personal-data

Market Condition	Open Standards and Interoperability	The connectivity of humans and machines relies on the seamless **interoperability** among disparate information technology systems. Interoperability means a computer program can communicate and exchange information with other computer programs and that both programs can use that information. Widespread interoperability across varying systems can only be achieved through reliance upon **standardtechnology** interfaces that establish clear rules for communicating.[21]

Table 9 Framework Conditions

[21] http://www.ecis.eu/open-standards/

3.2.2 Guiding Questions for the Demand Side

The overall objective is to analyse the strategic option of your demand side.

3.2.2.1 Adoption Chain Partner

For the analysis of adoption chain partners, we rely on (Adner 2006) who introduces a three step approach

Step1: Identification of partners and their contribution

- Which adoption chain partner, such as intermediaries, platform owner, ecosystem stakeholders, are required for ensuring market adoption of our solution?

Step 2: Adoption Chain Risk Assessment

- How important is the support / engagement of the co-adoption partner for the success of our innovation?
- What are potential risks and effects that we need to keep in mind?

Step 3: Benefit, Cost and Risk Calculation per Partner

- What is the role and interest of each identified adoption chain partner?
- What are his or her benefits as well as costs for adopting the innovation?
- What are potential risks and effects that need to be kept in mind?

3.2.2.2 Value Network Strategy

The Value network strategy is analysed in four steps:

Step 1: Is your innovation based on a platform business?

- Understanding the **various platform users** involved:
  - End User: Which actor can make decisions based on the generated output? Which main tasks are supported by the output? Which main benefits are provided by the user? Is your innovative offering solving the problem? Would users be willing to pay for the solution?
  - Complementary User Group: In case of limited willingness of end-user to pay (enough) for a solution, which other actors could be willing to pay for the solution? What is their role? Why would they be interested to engage?
- Understanding the **various currencies**:
  - In case of limited willingness of end-user to pay (enough) for a solution, in which ways can you benefit from them using your offering (e.g. generating valuable data assets through usage of the offering, making advertisement, e.g.)
- **Value unit** of your platform
  - What is the main item exchanged among platform users?

Step 2: What is your strategic role in the ecosystem?

- Which strategic role do you envision to play in the ecosystem (Niche Player, Keystone Player, Dominant Player or Smart Data Service Business?
- What do you need to do to get into this role?
- Why is this a promising strategy for you?

Step 3: How to harness network effects?

- On which level (data, infrastructure, marketplace) can your data-driven innovation benefit from network effects?
- What are concrete strategies to leverage networks effects? Some examples
  - on data level: data interoperability, data linkage, semantic models and knowledge graphs, customer locality programs, data portability, Incentives for data sharing, standardized data models, etc.
  - on infrastructure level: Open APIs, reference architectures, open standards,
  - On marketplace level: e.g. ensuring balanced growth of consumer and producer by freemium or similar incentivizing strategies?

Step 4: Roadmap and combination of revenue models?

- In order to address Step 1 to 3, which revenues models will you select for which user segment for what purpose? How will this change over time?
- Product and Price versioning: What bundles / combinations of offerings are you offering to each user segment? Can you bundle or unbundle the offering? Can you align the price/revenue model with the user needs?
- Competitiveness of each Value Proposition: Why is your offering better /more competitive than others offering on the market?

3.2.2.3 Framework Conditions

Which framework condition will impact my data-driven innovation?

- Which framework conditions are of relevance for the success of my envisioned data-driven innovation?
- How can you influence relevant framework conditions?
- How can you overcome barriers driven by framework conditions?

4 Engagement with SMEs and start-ups

We are aiming to use the DemoX methodology in the future for engaging with start-ups and SMEs to explore their business opportunities and investment decisions. A first interview with the data-driven start-up Dawex, demonstrates the practical value of the DemoX approach.

4.1 Interview with DAWEX

Dawex is a global data marketplace that enables you to share, monetize, and acquire data without intermediaries. It is designed for all types of private companies and public-sector organizations. They have built a unique trusted third-party platform that combines all the steps and functions for carrying out data transactions in proper form. Their goal is to contribute to building the data economy by making data exchange accessible and secure for all economic stakeholders.[22]

Value Proposition

Question: What is your value proposition and addressed customer segment?

- *Our vision is to become the global data exchange platform*
- *Our value proposition: global data exchange – we enable companies to meet and then to buy and sell data directly, we are not a broker (who sells and buys goods or assets for others)*

Question: How does your value proposition look like in concrete?

We help both partners to do the deal

- *Technical transaction to transfer the data*
- *Payment, i.e. we transfer the money*
- *Support to build the licensing model*
- *In case of personal data, we help to be compliant regarding the local rules and regulations*

Question: we are interested in your customer segment. Do you also offer closed market places, e.g. for data inside companies?

Actually, this is what we call Dawex for Private Data Exchange - in a private mode.

Data

Question: How does a typical data set look like traded by DAWEX?

[22] https://www.crunchbase.com/organization/dawex

I will give you 2 examples - there is nothing new on the market. There is data on potential car buyers for instance, as personal data, or when we think of industrial data think for instance of road network data coming from sensors. The information coming from sensors on the motorway as technical data.. tomorrow it will be IoT. Whatever the data, you can do a deal on Dawex - we don't care about the data - we do case about potential markets - about buyers and vendors. What the data is is not the principal matter for us?

Technology / Data Value Chain

Questions: What technology are you providing / using?

We are providing match-making algorithm based on the metadata describing the data sets, metadata description are provided by the data seller.

Value Network Strategy and Partner

Question: Is it right to say that your value proposition seems to be based on Multi-sided Platform Business Model and the value unit is data exchange (in comparison to stays at AIR BNB or Tweets at LinkedIn, travels at UBER or Videos at YouTube).

The aim of Dawex, is to become a global data exchange worldwide, the famous one, to be top of mind when it comes to data. In terms of data, it can be industrial data, personal data. Today we can handle both of it, because we have tuned the tool that was made for it. We can also deal with IoT data - we just launched a partnership with MNUBO (www.mnubo.com) who are an IoT company for analytics, mainly. SO basically, with Dawex you can do data transactions with any kind of data, any kind of company, anywhere in the world - we have around 15000 companies who on-boarded in the US and the EU - so the platform is operational, its on the market, its open. So, it is not a project, it is a company and let's say that the traction is very good today.

Question: And how are you making money, i.e. what are the revenue models you are using?

For the global data exchange we charge 5% of the data exchange as commission fee

For internal data exchange platform (e.g. inside a large corporate such as Siemens) we sell a license reflecting the number of features wanted, we make sure that this price is below

Ecosystem and Partner

Question: How do ensure the ecosystem building around your platform, i.e. how do you ensure balanced growth of supplier and consumer of data? Do you focus on a particular user segment to generate momentum?

- *No the strategy was clearly to be agnostic*
- *Speed is essential: on-board as many companies as possible, in order to avoid that companies get tired to wait. We expect that this is a winner-takes it model, in order to secure our today's and future position we require to be fast.*

- *We have dedicated services that support the on-boarding of new companies, i.e. we help to do the metadata description (Similar to AIRBND support apartment provider by sending and paying professional photographers)*
- *Active hunting for buyers in case of large data vendors or vice versa active hunting for sellers in case of a potential data vendor*

Question: How do you deal with balanced growth?

We are monitoring every single thing that is happening in the marketplace

Framework Conditions

Question: How is regulation and standards relevant for you?

This is part of our value proposition…we help companies to exchange data compliant to local rules and regulation.

Question: How will you business offering change when global regulations, such as the new GDPR, will be available?

The law is very strong. In general, this is very good for the market. When you want to be on this market you have to apply those rules, it is simple as that

Question: Which framework condition are supporting or hindering your business?

For our business scalability is key, in particular as scalability and speed are linked. Any framework conditions, standards that support scalability are of high relevance. An ideal ecosystem relies on free flow of data, and we rely on a healthy ecosystem, thus any framework conditions supporting the free flow of data are of high relevance

5 Conclusion and Next Steps

This report provides first update of the BDVe task aimed at documenting and reviewing emerging business opportunities. We reviewed literature and practice on existing business modelling approaches, looking at how well they are suited to capture the dynamics of data-driven propositions. Drawing on the main findings we designed and tested an updated approach for analysing data-driven business opportunities by combining features from existing approaches but with more emphasis on the data component as well as meta-analysis of demand and supply trends. While aiming to remain theoretically sound the approach also preserves the strength of the Business Model Canvas in that it is relatively easy to apply in workshop setting for instance when designing or testing business propositions from startups and SMEs.

In the next stage the model will be tested in a wider ranges of settings and contexts (verticals, horizontals, mixed) and fine-tuned based on the response from startups and corporates. The testing will be done in Business Model Canvas style webinars and workshops as well as through interviews and storytelling based on collected case studies. In addition we plan to evaluate our developed approach along the conducted teaching session as well as accomplish a quantitative analysis of a representative set of data-driven start-ups to identify promising patterns to guide future investment decisions.

To engage with the start-up community the sessions will be advertised and moderated through a portal integrated with the no. 1 start-up community in Europe (and the US) F6S[23]. Here we will establish links between the F6S platform and the BDV Portal. In addition we plan to conduct webinars with European stakeholders to guide them in future investment decisions.

[23] www.f6s.com

6 References

Adner, R., 2006. Match Your Innovation Strategy to Your Innovation Ecosystem Match Your Innovation Strategy to Your Innovation Ecosystem. *Harvard business review*, 84(4), pp.98–107.

Anderson, C., 2009. *Free*, Online Book.

Ardichvili, A. et al., 2003. A theory of entrepreneurial opportunity identification and development. *Journal of Business Venturing*, 18(1), pp.105–123.

Attenberger, Y., 2016. *Business Models and Big Data*, Master Thesis at the Munich School of Management, LMU Munich.

Baecker, D., 2007. *Studien zur nächsten Gesellschaft* 1. Aufl., Frankfurt am Main: Suhrkamp.

Cavanillas, J.M., Curry, E. & Wahlster, W., 2016. *New Horizons for a Data-Driven Economy: A Roadmap for Usage and Exploitation of Big Data in Europe*,

Choudary, S.P., 2015. *Platform Scale - How a new breed of startups is building large empires with minimal investment*, Platform Thinking Labs Pte Ltx.

Curry, E., 2016. The Big Data Value Chain: Definitions, Concepts, and Theoretical Approaches. In *New Horizons for a Data-Driven Economy: A Roadmap for Usage and Exploitation of Big Data in Europe*.

Davis, L.S. et al., 2014. Learning to Segment Object Candidates. *IEEE Conference on Computer Vision and Pattern Recognition*, (March 2015), pp.1–10.

Gassmann, O., Frankenberger, K. & Csik, M., 2014. Revolutionizing the Business Model - St. Gallen Business Model Navigator. *Management of the Fuzzy Front End of Innovation*, 18(3), pp.89–97.

Hagel, J. et al., 2015. Patterns of disruption - Anticipating disruptive strategies in a world of unicorns, black swans and exponentials. *Deloitte University Press*, p.28`.

Hartmann, P.M. et al., 2014. Big Data for Big Business? A Taxonomy of Data-driven BusinessModels used by Start-up Firms. *Cambridge Service Alliance Whitepaper*, pp.1–29. Available at: http://cambridgeservicealliance.blogspot.co.uk/2014/04/big-data-for-big-business_3.html.

Iansiti, M. & Levien, R., 2004. *The keystone advantage*,

IDC & OpenEvidece, 2017. The European Data Market Study: Final Report. , p.275. Available at: http://datalandscape.eu/study-reports/european-data-market-study-final-report.

Johnson, M.W., Christensen, C.M. & Kagermann, H., 2008. Reinventing your business model. *Harvard Business Review*, 86(12).

Nooren, P. et al., 2014. MEDIA-INTERNET-TELECOM VALUE WEB INTRODUCING THE DAMIAN METHOD FOR SYSTEMATIC ANALYSIS OF THE INTERDEPENDENCIES BETWEEN MEDIA-INTERNET-TELECOM VALUE WEB. *TNO Report*.

OECD, 2015. Data-Driven Innovation: Big Data for Growth and Well-Being. *OECD Digital Economy Papers*, p.456. Available at: http://www.oecd-ilibrary.org/content/book/9789264229358-en.

Osterwalder, A. et al., 2014. *Value proposition design*,

Osterwalder, A. & Pigneur, Y., 2010. *Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers*, Toronto: Self Published.

Parker, G.G., Van Alstyne, M.W. & Choudary, S.P., 2016. Platform revolution: How Networked Markets Are Transforming the Economy--And How to Make Them Work for You. *W.W. Norton & Company*, p.256.

Shapiro, C. & Varian, H.R., 1999. *Information rules: A Stragegic Guide to The Network Economy*,

Timmons, J.A. & Spinelli, S., 2007. *New Venture Creation: Entrepreneurship for the 21st Century*, Available at: http://www.amazon.com/dp/0071276327.

Zillner, S. et al. eds., 2017. *Big Data Value Strategic Research and Innovation Agenda*.