D2.6 SECOND INTERIM REPORT

THE EUROPEAN DATA MARKET MONITORING TOOL: KEY FACTS & FIGURES, FIRST POLICY CONCLUSIONS, DATA LANDSCAPE AND QUANTIFIED STORIES

UPDATE OF THE EUROPEAN DATA MARKET STUDY

SMART 2016/0063

IDC Italia srl (Milan, IT)
The Lisbon Council (Brussels, BE)

Prepared for:

Katalin IMREI

Policy Officer

European Commission – DG CONNECT

Unit G1 – Data Policy & Innovation

28th June 2019

Author(s)	Giorgio Micheletti, Cristina Pepato (IDC)
Deliverable	D2.6 Second Interim Report
Date of delivery	28th June 2019
Version	1.0
Addressee officer	Katalin IMREI Policy Officer European Commission, DG CONNECT Unit G1 — Data Policy and Innovation EUFO 1/178, L-2557 Luxembourg/Gasperich katalin.imrei@ec.europa.eu
Contract ref.	N- 30-CE-0835309/00-96

Contents

Table of Figures

Table of Tables

EXECUTIVE SUMMARY

This is the Second Interim Study Report (Deliverable D2.6) of the Update of the European Data Market Study (SMART 2016/0063), entrusted in 2016 to IDC and the Lisbon Council. The present document brings together the research results and the activities carried out by the contractors under:

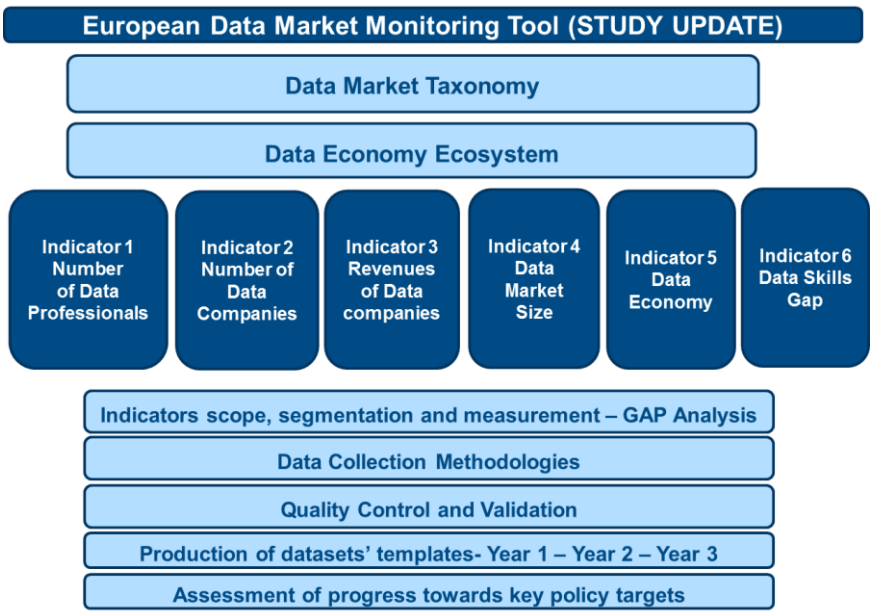
- **The Second Report on Facts & Figures (D2.4)** extending the measurement of the European Data Market Monitoring Tool by presenting data for the years 2016-2018 and forecasts to the year 2025 under three alternative scenarios;
- **The Second Report on Policy Conclusions (D2.5)** measuring the progress of European policies towards the objective of maximising the growth of the Data Economy as measured by the European Data Market Monitoring Tool;
- The key messages from the **quantified stories (D3.3, D3.4 and D3.5)** produced by the study team and focusing on the operational, organizational and/or economic benefits generated by the use of data-driven technologies with a special focus on Big Data and Artificial Intelligence (AI) solutions;
- **The Second Data Landscape Report (January 2019 Review – D4.2)** providing an overview of the EU Data Landscape and offering and up-to-date zoom into the database of data market companies in Europe.

Quantifying the European Data Market – Key Facts & Figures

The European Data Market Monitoring Tool

In line with the results presented in the original European Data Market study (SMART 2013/0063) in February 2017, the First Report on Facts & Figures (D2.1) of February 2018 and the Second Report on Facts & Figures (D2.4) of March 2019, the measured indicators are organised around a modular and flexible structure – the European Data Market Monitoring Tool. The updated European Data Market Monitoring Tool designed by IDC is shown in the Figure below.

Figure 1: The Updated EDM Monitoring Tool



The EU Data Market and Data Economy in 2018

The value of the **Data Economy**, which measures the overall impacts of the Data Market on the economy as a whole, is to exceed the threshold of 300 Billion Euro in 2018 for EU28, with a growth of nearly 12% over the previous year. Brexit uncertainties, however, play an important role in affecting the results for the EU28, which shows lower growth in 2018 compared to the EU27 (12.6%). The positive trend in the growth of the Data Economy is confirmed by the EU28 Data Market value in 2018, which is displaying a growth rate well above the one exhibited by the total IT spending, at 9.7% year-on-year, reaching 71.5 Billion Euro. This positive dynamic is registered in all Member States, even though the pace of growth varies.

As far as supply and demand are concerned, **data suppliers** are estimated at more than 283,000 units in the EU28 for 2018, exhibiting a year-on-year growth of 4.2%. **Data users**, instead, are projected to grow at 3.4% in 2018, amounting to more than 715,000 units. Following increasing growth rates over the prior four years, these figures show a picture of consolidation of data companies in the EU. **Revenues generated by data suppliers** have increased by 12% to reach 77 Billion Euro in the EU28 with the U.K., still in the leading position, Germany, France, Italy, the Netherlands and Spain showing the highest share of data revenues per country - together accounting for more than three quarters of data revenues in the European Union.

According to the latest estimates, the number of **data professionals** in the EU28 reached 7.2 million in 2018, corresponding to 3.4% of the total workforce, with an increase of 8% over the previous year. However, in 2018, the EDM Monitoring Tool continues to register an imbalance between the **demand and the supply of data skills** in Europe as the estimated gap grew by 18% reaching approximately 571,000 unfilled positions in the EU28, corresponding to 7.2% of total demand. This gap is not expected to decrease over the next few years, pointing to the need for policy intervention to further develop data skills in Europe.

The EU Data Market and Data Economy in 2025

The Update of the European Data Market Study also produced key facts & figures for the year 2025 according to three alternative evolution paths of the European Data Market and Economy, driven by different macroeconomic and framework conditions, highlighting the critical turning points to be faced in the next years by governments, businesses and social actors. The scenarios at 2025 continue to take as a reference point the initial scenarios developed for the year 2020. While the 2020 scenarios were mainly differentiated by economic drivers (different demand-supply dynamics), the 2025 scenarios continue to be shaped by a combination of economic and social drivers, focused on the interaction of two main focal issues (or evolution paths):

- The high or low pace of diffusion of data-driven innovation, driven by demand-supply dynamics, and its impact on economic growth;
- The high or low concentration of power in the access, control and exploitation of data assets, that is the social model of data governance.

The Data Economy scenarios are positioned at the intersection of these two main focal issues and take the following shape:

- The **Baseline scenario** is characterised by a healthy growth of data innovation, a moderate concentration of power by dominant data owners with a data governance model protecting personal data rights, and an uneven but rather wide distribution of data innovation benefits in the society;

- The **High Growth scenario** (Data-driven reality) is characterised by a high level of data innovation, low data power concentration, an open and transparent data governance model with high data sharing, and a wide distribution of the benefits of data innovation in the society;
- The **Challenge scenario** (Digital Maze) is characterised by a low level of data innovation, a moderate level of data power concentration due to digital markets fragmentation, and an uneven distribution of data innovation benefits in the society.

In the **Baseline scenario** the EU GDP cumulative growth average in the period 2018-2025 (+1.7%) will sustain the investments in the digital economy and consumer willingness to spend. As a result, the Data Market is forecast to reach 83 billion Euro in the EU27 and 106 billion in the EU28, with a compound annual growth rate of 6.5% and 6.4% respectively between 2020 and 2025. The Data Economy will grow faster than the Data Market, thanks to a positive multiplier impact of data innovation on the economy, reaching a value of 555 billion Euro in the EU27 and 680 billion Euro in the EU28, doubling its incidence on EU GDP to 4% and 4.2% respectively, compared to 2.2% for EU27 and 2.4% for EU28 in 2017. Enterprises will add more than 3 million data professionals' positions between 2020 and 2025. However, this will increase the potential data professionals' skills gap to approximately 925,000 unfilled positions in the EU27 and more than 1 million unfilled position in the EU28, corresponding respectively to 10% and 9.2% of total data skills demand. The lack of skills may become a bottleneck for some enterprises or regions, as data talent "wars" will likely develop for the most skilled professionals.

In the High Growth scenario at 2025, the EU GDP compound annual growth rate in the period 2020-2025 (+2.2%) will be 2.5 times higher than in the Challenge scenario and 1.5 times higher than in the Baseline scenario. This will accelerate the investments in the digital economy and consumer willingness to spend. In the European Union public and private investments will accelerate in Artificial Intelligence, advanced robotics, automation as well as new skills. As a result, the Data Market is forecast to reach 107 billion Euro in the EU27 and 142 billion Euro in the EU28, with a compound annual growth rate of 12% and 12.7% respectively between 2025 and 2020. The Data Economy will reach a value of 829 billion Euro in the EU27 and 1,054 billion Euro in the EU28, with an incidence on EU GDP of 5.8% and 6.3% respectively, compared to 2.4% and 2.6% in 2018. Enterprises will add more than 4 million data professionals' positions between 2020 and 2025 (compared to 3 million in the previous scenario). However, the potential data professionals' skills gap will grow exponentially to 1.5 million unfilled positions in the EU27 and 1.7 million unfilled positions in the EU28, corresponding respectively to 14% and 13% of total data skills demand. This risk will need to be managed in advance or the lack of skills will become a serious constraint for data-driven companies and data suppliers.

In the Challenge scenario, the EU GDP compound annual growth rate in the period 2020-2025 will be only 0.9%, that is substantially lower than the other scenarios. As a result, the Data Market is forecast to approach 72 billion Euro in the EU27 and 93 billion in the EU28, respectively marking a compound annual growth rate of 3.6% between 2020 and 2025. In the same context, the Data Economy will reach a value of 444 billion Euro in the EU27 and 546 billion Euro in the EU28 with an incidence on GDP of 3.3% and 3.5% respectively, compared to 2.4% for EU27 and 2.6% for EU28 in 2018, or to 4% and 4.2% respectively in the Baseline scenario 2025.

The number of data professionals will still increase to 8.5 million in the EU27 and 10.3 million in the EU28 in 2025, adding 2 million data professionals' positions compared to 2020. We estimate a

potential data skills gap of approximately 775,000 unfilled positions in the EU27 and 829,000 unfilled positions in the EU28 in 2025, corresponding respectively to 9% and 8% of total demand, as demand will still grow faster than supply. The lower supply will be due mainly to lower market entries from other careers and less upskilling-retraining initiatives, because of the lower attractiveness of the Data Market. The uneven diffusion of data innovation will result in a mismatch between demand and supply by geographical area across the Union, with unemployment in some regions and unsatisfied demand in others.

The EU Data Market and the International Indicators

The latest estimates of the European Data Market Monitoring Tool confirm the picture drawn by the previous measurements of the international indicators - the EU28 is the second-largest Data Economy worldwide after the U.S., ahead of Brazil and Japan (the other two EU international partners considered by the European Data Market Study).

The U.S. continue to show the highest impact of the Data Economy on their GDP – 1.17% in 2018, up 13.4% with respect to 2017. While not leading in absolute values, however, Europe emerges as the most dynamic region with a sustained and unsurpassed impacts' growth of more than 9% year-on-year 2017 at the level of the EU27. In terms of size and growth, the value of its Data Market is second only to the U.S.; more interestingly, the impact that this market generates on the economy as a whole (the "Data Economy") has become more and more visible over the past few years thus rapidly catching up the gap with the American economy.

Regarding the number of data suppliers, though, Europe's growth slows to the lowest of the group, as the U.S. move ahead, although Europe still presses the U.S. in terms of the total number of data suppliers with nearly as many. Europe also easily out-competes Japan though with two companies for every single company in Japan.

The four infographics at the end of this Executive Summary provide a comprehensive overview of the main facts and figures stemming from the Update of the European Data Market Study so far.

Describing the Data Market – The Quantified Stories

The quantified stories were the result of a mixed effort entailing both secondary and primary research activities. Extensive secondary research on available public sources, specialised press and academic literature was undertaken to obtain an actionable and up-to-date understanding of the operational, organizational and/or economic benefits generated by the use of data-driven technologies with a special focus on Big Data and Artificial Intelligence (AI) solutions. Three stories were produced by the study team between August 2018 and May 2019. While the first story investigated the relationship between the use of data and the means to achieve additional revenues in a business-to-business environment, the second and the third stories focused on the role of data in nurturing and expanding the benefits brought about by AI implementations across a number of different industries and, specifically, on the energy and utilities sector.

More specifically, the first story presented in this report (Story 3 of the EDM study) highlighted that, while still in its infancy, data monetization already constitutes a powerful means to generate additional revenues by using data to add new services to existing offerings, developing new business models, and even directly selling data-based products, services or utilities. Through a series of case-studies featuring real-life examples of AI implementations across multiple sectors, the first story unveiled a rather complicated situation in Europe in terms of data monetization. The key findings stemming from the analysis are as follows:

- A clear-cut form of data monetization is hard to find.
- Benefits such operational efficiency, cost optimization, and enhanced quality seem to prevail vis-à-vis data monetization per se.
- Platforms are the preferred means to perform and enable data monetization.
- Medium to large companies are at the forefront of data monetization.
- Wide differences in digital maturity, uncertainty surrounding data ownership, and lack of data skills hamper the development of data monetization.

To better understand how European companies are approaching and implementing AI in a variety of different sectors, the second story presented in this report (Story 4 of the EDM Study) featured extensive desk research across a multitude of publicly available sources and IDC existing research material, and produced five case-studies showcasing how AI solutions currently adopted in Europe to advance an organization's predictive capabilities allowing for a better interpretation of customer needs and improved forecast accuracy.

The analysis found that AI is embraced to extend the quality of products and services, as well as to obtain overall business benefits in sales & marketing campaigns and customer loyalty programmes. The case studies presented in this research highlight a number of immediate benefits obtained by the organizations adopting AI and Big Data technologies, mainly linked to ease of use and a simplified technology, and to more advanced predictive capabilities.

The third story presented in this report (Story 5 of the EDM Study) analysed how European utilities are increasingly turning to data-driven technologies to excel in operations, improve customer service satisfaction and create new revenue streams and business models. Indeed, Big Data technologies, smart meter solutions and the increasing deployment of the Internet of Things (IoT), are fuelling an operational and technological revolution in European utilities and leading to the rapid introduction of Artificial Intelligence (AI). In turn, AI is starting to emerge as the technology game changer for the industry, helping utilities make operation more cost effective while ensuring optimal utilization of infrastructure and resources to balance supply and demand safety and reliability.

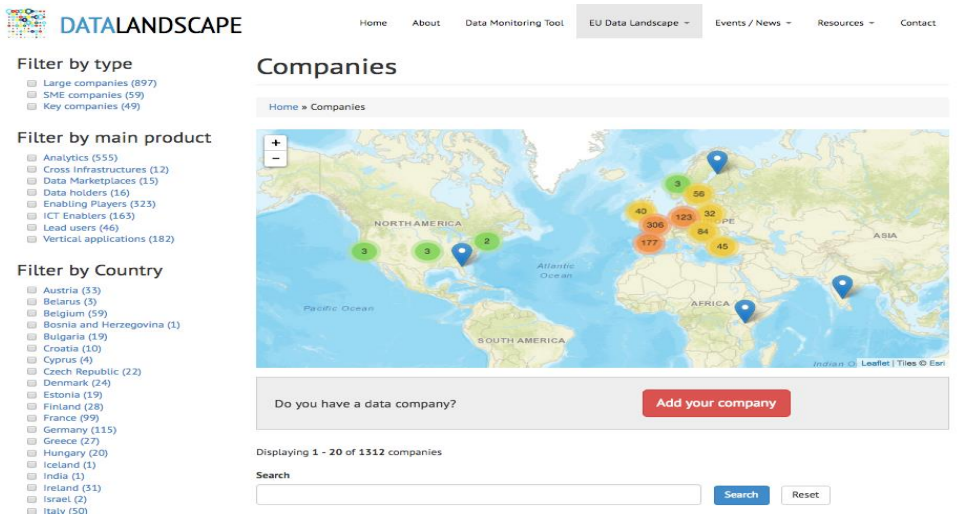
The desk research and in-depth interviews carried out across the six case studies presented in this research confirm that AI is being successfully employed across a number of European utilities to improve or transform the organisation's operating model, enhance the customer experience, and create new revenue streams and business models.

Mapping the Data Market – The Data Landscape and the Data Market Monitoring Tool

The Second EU Data Landscape Report (D4.2) provides an overview of the EU Data Landscape database revision as of January 2019.

With a total of 1459 companies and coverage of 41 countries (European Union-28, Belarus, Bosnia and Herzegovina, Georgia, Iceland, Israel, Kenya, Moldova, Norway, Serbia, Switzerland, Turkey, Ukraine and the United States), the database has grown by 16% from 2018 to 2019 with the addition of 203 new companies. Out of the new companies, 165 were identified as Key Data Landscape companies, offering a comprehensive overview of the most important data companies in Europe.

Figure 2: Database of Data Landscape companies (datalandscape.eu)

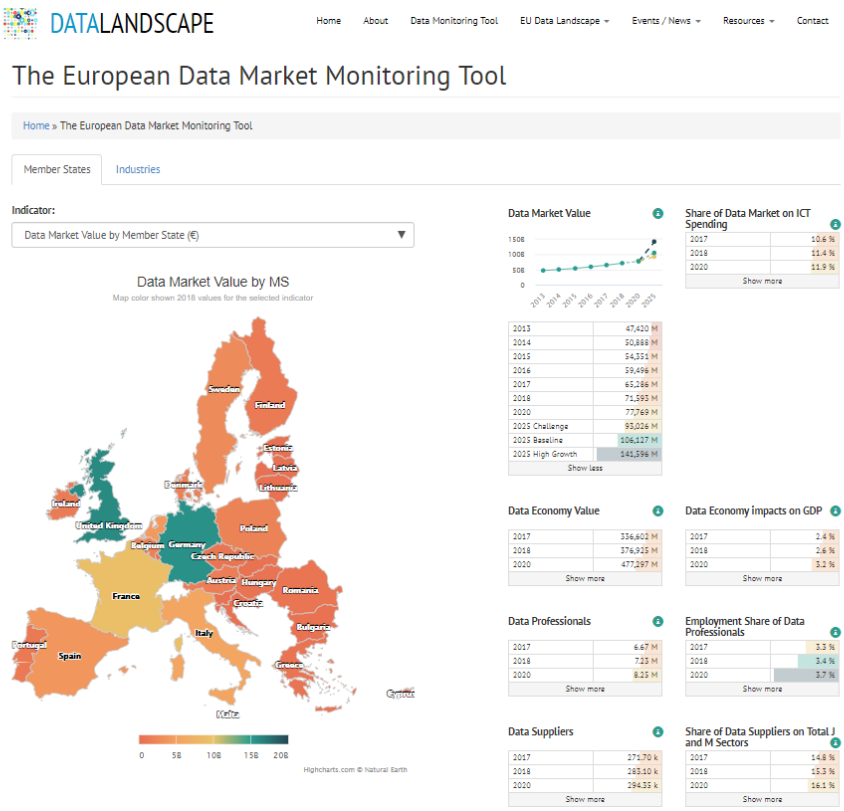


Source: http://datalandscape.eu/companies

Together with the extension and the validation of the EU Data Landscape database, the study team performed an update of the European Market Monitoring Tool launched on www.datalandscape.eu in April 2018.

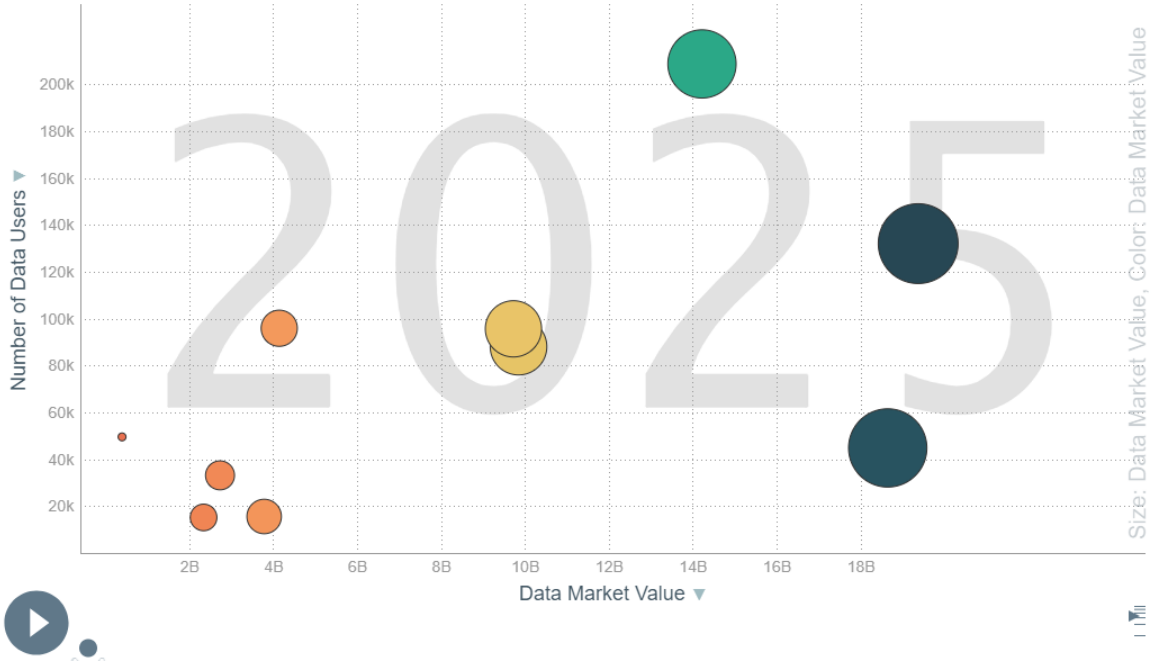
The update, based on the Dataset accompanying the Second Report on Facts and Figures (D2.4), extends the scope of the EDM Monitoring Tool to include the key facts and figures measured for the year 2018 at the level of the total EU28 and for all 28 EU Member States, when available and applicable. Industry-specific views were also offered with indicators provided by industry sector when possible. The figures presented for the years 2020 and 2025 (three scenarios) were also updated to reflect the results obtained from the second round of measurement of the EDM indicators.

Figure 3: The European Data Market Monitoring Tool



Source: http://datalandscape.eu/european-data-market-monitoring-tool-2018

Figure 4: The European Data Market Monitoring Tool – Bubble chart



Source: http://datalandscape.eu/european-data-market-monitoring-tool-2018

Acting Upon the Data Market – The Role of Policy

Maximising the growth of the Data Market and Data Economy is a key objective of the EU policy action and specifically of the Digital Single Market (DSM) Strategy, deployed through the initiatives launched in 2016-18, such as the Digitizing European Industry initiative, the European Cloud initiative and the AI Action Plan. The need for Europe to capture the digital opportunity is expected to remain at the forefront of European strategies and further contribute to the development of data-driven innovation.

Europe's Data Market and Data Economy Evolution: Policy and the Three Scenarios

The year 2018 has seen all the indicators measured by the EDM tool in a positive growth dynamic from 2017, as the European economy continues a positive cycle of development. This positive dynamic is registered in all Member States, even though the pace of growth varies, with small economies pushing forward their digitisation strategies, such as Slovakia, Lithuania, Slovenia and Latvia, thus exhibiting significantly higher growth rates than the EU average.

The way Europe's Data Market and Data Economy may evolve in the next years remains open: The 2025 scenarios developed for Update of the European Data Market Study (SMART 2016/0063) present potential evolution paths of the European Data Market and Economy and are built around the intersection of two main focal issues:

- The Data Market's pace of growth;
- The potential evolution of the model of data governance.

In the Baseline scenario policy is expected to reach only a partial success. While GDPR has set itself as a worldwide standard for privacy protection, gaining attention and being emulated also in the U.S., adjustments and revisions linked to implementations hurdles are also expected to come. The so called GAFA (Google, Facebook, Amazon, Apple) will maintain their dominant position a regards personal data flows but will have to become more transparent and guarantee users more control on how their personal data is used and shared. At the same time, the entry into force of the FFoD Regulation will be essential in order to unlock the exploitation of European datasets and enable new data-driven processes such as machine learning. The completion of the Digital Single Market (DSM) is expected to progress and possibly be completed only at the end of the forecast period. In this scenario, European investments in digital innovation will increase strongly through the Horizon Europe and the new Digital Europe Programme, with support by industry through the Public-Private Partnerships and their collaborations. A more relevant policy role is foreseen under the High Growth scenario. The completion of the Digital Single Market within the forecast period is a key success factor: this requires rapid and successful implementation of the GDPR and the Free-Flow of non-personal data initiative, with true liberalisation of data flows across Europe. R&D investment at EU and national level must be particularly effective and the Network of Excellence Centres must succeed in becoming innovation beacons in every region as well as increasing the supply of scarce data skills. The Challenge scenario is driven as much by the failure of the Digital Single Market and of innovation investments than by global economic trends. The risk of insufficient digital capacities and networking powering data innovation across Europe would create a disadvantage for the European industry and a growth gap with other world regions, particularly Asia-pacific.

Value from Data, the new Priority of Data Users

According to our latest estimates, the European Data Market has increased by over 50% from 2014 to 2018 (from €47B to €72B in the EU28), which represents a remarkable increase. But even more remarkable has been the change in awareness about the value of data and the potential of data-

driven innovation to increase growth and welfare. Collecting, analysing and exploiting data is understood as the enabling condition of digital transformation, as technologies such as the Internet of Things, Virtual/Augmented Reality, Cybersecurity, Robotics, Additive manufacturing (3D printing) rely on the collection and manipulation of data to be implemented.

In this framework, the connection between Big Data and Artificial Intelligence is establishing itself as the real game changer. The European Union has recognised the relevance of the forthcoming AI revolution by launching an AI Communication and Action Plan, coordinating efforts and increasing investments to make sure that Europe keeps pace with international competition. At the same time, Europe has chosen to deal also with the potential risks of Artificial Intelligence, aiming for a human-centered AI, respecting the principles of privacy protection and basic human rights.

Nevertheless the completion of the mandate of the current European Commission and Parliament, in 2019, the need for Europe to capture the digital opportunity will remain at the forefront of European strategies. Next year, European Parliament and the Council will finalise the new Multi-Annual Financial Framework 2021-2027 proposed by the Commission making decisions which will influence the size and directions of the investments in digital technologies and policies. In this scenario, market intelligence is needed more than ever to inform and support strategic decisions.

The EU Data Policy and the International Dimension

The international indicators emerging from the second round of measurements of the European Data Market Monitoring Tool continue to highlight the prominence of the U.S. in terms of size of the Data Market and the Data Economy, even if Europe has undertaken significant steps in 2018 towards a more coordinated approach and a progressive removal of the barriers identified by the Digital Single Market Strategy and continues to represent the second-largest Data Economy worldwide after the U.S. Europe, however, still lags behind in terms of digital infrastructure and number of data professionals. Not only does it suffer from high levels of fragmentation when it comes to the use of digital technologies across companies, but it is also affected by the lack of the big tech corporations, which represent two of the main reasons linked to Europe's relative delay. Filling this gap would be essential to increase Europe's competitiveness and for the future of work in the EU, which requires a digital upskilling of the workforce.

As regards the progressive removal of the challenges identified by the DSM Strategy, noteworthy policy measures in 2018 in this respect have been the entry into force of the General Data Protection Regulation, the Regulation on the free flow of non-personal data, as well as the Regulation addressing unjustified geo-blocking. In 2019 other important steps have been made through the European Parliament's and Council of the European Union's approval of the new Copyright Directive during last March and April, while intensified efforts towards the adoption of an ePrivacy Regulation, necessary to complete the EU's framework for data protection and confidentiality of communications, will probably foster discussion in the next period. With the aim of filling the gap with other world regions, and in particular the U.S., the EU has also recently launched new policy initiatives in the field of AI.

Within this framework, Europe's efforts in setting standards not only as far as ethics principles are concerned, but also as regards data protection, privacy and fair competition, have led to greater attention towards these topics on an international level and can put the grounds for Europe to be a role model for other countries.

Indicator 1: Data Professionals

Workers who collect, store, manage, analyse, interpret, and visualise data as their primary or as a relevant part of their activity.

Million	2017	2018	Growth '18/'17	Share of data professionals on total EU employment, '18
	6.6	7.2	8.4%	3.4%

Indicator 6: Data Professionals Skills Gap

The indicator captures the potential gap between demand and supply of data skills in Europe.

Thousand	2017	2018	Growth '18/'17	Share of total skills demand, '18
	483	571	18%	7.2%

Indicator 2: Data Companies

Data suppliers

Data suppliers have as their main activity the production and delivery of digital data-related products, services, and technologies.

Thousand	2017	2018	Growth '18/'17	Companies Share as a % of total companies in ICT and Professional Services, '18
	271	283	4.2%	15%

Data users

Data users are organisations that generate, exploit collect and analyse digital data intensively and use what they learn to improve their business.

Thousand	2017	2018	Growth '18/'17	Companies Share as a % of total companies, '18
	691	715	3.4%	6.8%

Indicator 3: Data suppliers' revenues

The aggregated value of all the data-related products and services generated by EU Data suppliers companies.

€ Bn	2017	2018	Growth '18/'17
	69	77	12%

Indicator 4: Value of the Data Market

The marketplace where digital data is exchanged as "products" or "services" as a result of the elaboration of raw data.

€ Bn	2017	2018	Growth '18/'17
	65	71	9.7%

Indicator 5: Value of the Data Economy

The Data Economy measures the overall impacts of the data market on the economy as a whole.

€ Bn	2017	2018	Growth '18/'17	Share of EU GDP, '18
	336	377	12%	2.6%

Source: EDM Monitoring Tool, IDC 2019

European Commission

IDC ANALYZE THE FUTURE

theLisboncouncil think tank for the 21st century

Indicator 1: Data Professionals

Workers who collect, store, manage, analyse, interpret, and visualise data as their primary or as a relevant part of their activity.

Million | 2020: 6.5 | 2025: 9.3 | CAGR '25/'20: 7.3%

Indicator 6: Data Professionals Skills Gap

The indicator captures the potential gap between demand and supply of data skills in Europe.

Thousand | 2020: 603 | 2025: 925 | CAGR '25/'20: 9% | Share of total skills demand, '25: 10%

Indicator 2: Data Companies

Data suppliers

Data suppliers have as their main activity the production and delivery of digital data-related products, services, and technologies.

Thousand | 2020: 151 | 2025: 174 | CAGR '25/'20: 3% | Companies Share as a % of total companies in ICT and Professional Services, '25: 13.6%

Data users

Data users are organisations that generate, exploit collect and analyse digital data intensively and use what they learn to improve their business.

Thousand | 2020: 538 | 2025: 581 | CAGR '25/'20: 1.5% | Companies Share as a % of total companies, '25: 6.5%

Indicator 3: Data Suppliers' revenues

The aggregated value of all the data-related products and services generated by EU Data suppliers companies.

€ Bn | 2020: 70 | 2025: 97 | CAGR '25/'20: 6.8%

Indicator 4: Value of the Data Market

The marketplace where digital data is exchanged as "products" or "services" as a result of the elaboration of raw data.

€ Bn | 2020: 60 | 2025: 83 | CAGR '25/'20: 6.5%

Indicator 5: Value of the Data Economy

The indicator captures the potential gap between demand and supply of data skills in Europe.

€ Bn | 2020: 387 | 2025: 555 | CAGR '25/'20: 7.5% | Share of EU GDP, '25: 4%

Source: EDM Monitoring Tool, IDC 2019

Indicator 1: Data Professionals

Workers who collect, store, manage, analyse, interpret, and visualise data as their primary or as a relevant part of their activity.

Million	2020	2025	CAGR '25/'20
	8.3	11.3	6.6%

Indicator 6: Data Professionals Skills Gap

The indicator captures the potential gap between demand and supply of data skills in Europe.

Thousand	2020	2025	CAGR '25/'20	Share of total skills demand, '25
	641	1039	10%	9.2%

Indicator 2: Data Companies

Data suppliers

Data suppliers have as their main activity the production and delivery of digital data-related products, services, and technologies.

Thousand	2020	2025	CAGR '25/'20	Companies Share as a % of total companies in ICT and Professional Services, '25
	294	339	2.8%	17.6%

Data users

Data users are organisations that generate, exploit collect and analyse digital data intensively and use what they learn to improve their business.

Thousand	2020	2025	CAGR '25/'20	Companies Share as a % of total companies, '25
	721	779	1.5%	7.4%

Indicator 3: Data Suppliers' revenues

The aggregated value of all the data-related products and services generated by EU Data suppliers companies.

€ Bn	2020	2025	CAGR '25/'20
	90	126	7.1%

Indicator 4: Value of the Data Market

The marketplace where digital data is exchanged as "products" or "services" as a result of the elaboration of raw data.

€ Bn	2020	2025	CAGR '25/'20
	78	106	6.4%

Indicator 5: Value of the Data Economy

The indicator captures the potential gap between demand and supply of data skills in Europe.

€ Bn	2020	2025	CAGR '25/'20	Share of EU GDP, '25
	477	680	7.3%	4.2%

Source: EDM Monitoring Tool, IDC 2019

European Commission

IDC ANALYZE THE FUTURE

theLisboncouncil think tank for the 21st century

Indicator 1: Data Professionals

Workers who collect, store, manage, analyse, interpret, and visualise data as their primary or as a relevant part of their activity.

Million | 2020: 6.5 | 2025: 11 | CAGR '25/'20: 10.6%

Indicator 6: Data Professionals Skills Gap

The indicator captures the potential gap between demand and supply of data skills in Europe.

Thousand | 2020: 603 | 2025: 1,551 | CAGR '25/'20: 21% | Share of total skills demand, '25: 14.2%

Indicator 2: Data Companies

Data suppliers

Data suppliers have as their main activity the production and delivery of digital data-related products, services, and technologies.

Thousand | 2020: 151 | 2025: 194 | CAGR '25/'20: 5.2% | Companies Share as a % of total companies in ICT and Professional Services, '25: 15%

Data users

Data users are organisations that generate, exploit collect and analyse digital data intensively and use what they learn to improve their business.

Thousand | 2020: 538 | 2025: 624 | CAGR '25/'20: 3% | Companies Share as a % of total companies, '25: 7%

Indicator 3: Data Suppliers' revenues

The aggregated value of all the data-related products and services generated by EU Data suppliers companies.

€ Bn | 2020: 70 | 2025: 134 | CAGR '25/'20: 14%

Indicator 4: Value of the Data Market

The marketplace where digital data is exchanged as "products" or "services" as a result of the elaboration of raw data.

€ Bn | 2020: 60 | 2025: 107 | CAGR '25/'20: 12%

Indicator 5: Value of the Data Economy

The indicator captures the potential gap between demand and supply of data skills in Europe.

€ Bn | 2020: 387 | 2025: 829 | CAGR '25/'20: 16.4% | Share of EU GDP, '25: 5.8%

Source: EDM Monitoring Tool, IDC 2019

European Commission

IDC ANALYZE THE FUTURE

the Lisbon council think tank for the 21st century

Indicator 1: Data Professionals

Workers who collect, store, manage, analyse, interpret, and visualise data as their primary or as a relevant part of their activity.

Million	2020	2025	CAGR '25/'20
	8.3	13	9.8%

Indicator 6: Data Professionals Skills Gap

The indicator captures the potential gap between demand and supply of data skills in Europe.

Thousand	2020	2025	CAGR '25/'20	Share of total skills demand, '25
	641	1,693	21.4%	13%

Indicator 2: Data Companies

Data suppliers

Data suppliers have as their main activity the production and delivery of digital data-related products, services, and technologies.

Thousand	2020	2025	CAGR '25/'20	Companies Share as a % of total companies in ICT and Professional Services, '25
	294	389	5.7%	20%

Data users

Data users are organisations that generate, exploit collect and analyse digital data intensively and use what they learn to improve their business.

Thousand	2020	2025	CAGR '25/'20	Companies Share as a % of total companies, '25
	721	844	3.2%	8%

Indicator 3: Data Suppliers' revenues

The aggregated value of all the data-related products and services generated by EU Data suppliers companies.

€ Bn	2020	2025	CAGR '25/'20
	90	179	14,8%

Indicator 4: Value of the Data Market

The marketplace where digital data is exchanged as "products" or "services" as a result of the elaboration of raw data.

€ Bn	2020	2025	CAGR '25/'20
	78	142	24.3%

Indicator 5: Value of the Data Economy

The indicator captures the potential gap between demand and supply of data skills in Europe.

€ Bn	2020	2025	CAGR '25/'20	Share of EU GDP, '25
	477	1,054	17.2%	6.3%

Source: EDM Monitoring Tool, IDC 2019

Indicator 1: Data Professionals

Workers who collect, store, manage, analyse, interpret, and visualise data as their primary or as a relevant part of their activity.

Million | 2020: 6.5 | 2025: 8.5 | CAGR '25/'20: 5.5%

Indicator 6: Data Professionals Skills Gap

The indicator captures the potential gap between demand and supply of data skills in Europe.

Thousand | 2020: 603 | 2025: 775 | CAGR '25/'20: 5.2% | Share of total skills demand, '25: 9%

Indicator 2: Data Companies

Data suppliers

Data suppliers have as their main activity the production and delivery of digital data-related products, services, and technologies.

Thousand | 2020: 151 | 2025: 163 | CAGR '25/'20: 1.6% | Companies Share as a % of total companies in ICT and Professional Services, '25: 13%

Data users

Data users are organisations that generate, exploit collect and analyse digital data intensively and use what they learn to improve their business.

Thousand | 2020: 538 | 2025: 558 | CAGR '25/'20: 0.7% | Companies Share as a % of total companies, '25: 6.3%

Indicator 3: Data Suppliers' revenues

The aggregated value of all the data-related products and services generated by EU Data suppliers companies.

€ Bn | 2020: 70 | 2025: 77 | CAGR '25/'20: 2%

Indicator 4: Value of the Data Market

The marketplace where digital data is exchanged as "products" or "services" as a result of the elaboration of raw data.

€ Bn | 2020: 60 | 2025: 72 | CAGR '25/'20: 3.6%

Indicator 5: Value of the Data Economy

The indicator captures the potential gap between demand and supply of data skills in Europe.

€ Bn | 2020: 387 | 2025: 444 | CAGR '25/'20: 2.8% | Share of EU GDP, '25: 3.3%

Source: EDM Monitoring Tool, IDC 2019

Indicator 1: Data Professionals

Workers who collect, store, manage, analyse, interpret, and visualise data as their primary or as a relevant part of their activity.

Million | 2020: 8.3 | 2025: 10.3 | CAGR '25/'20: 4.6%

Indicator 6: Data Professionals Skills Gap

The indicator captures the potential gap between demand and supply of data skills in Europe.

Thousand | 2020: 641 | 2025: 829 | CAGR '25/'20: 5.3% | Share of total skills demand, '25: 8%

Indicator 2: Data Companies

Data suppliers

Data suppliers have as their main activity the production and delivery of digital data-related products, services, and technologies.

Thousand | 2020: 294 | 2025: 321 | CAGR '25/'20: 1.7% | Companies Share as a % of total companies in ICT and Professional Services, '25: 17%

Data users

Data users are organisations that generate, exploit collect and analyse digital data intensively and use what they learn to improve their business.

Thousand | 2020: 721 | 2025: 751 | CAGR '25/'20: 0.8% | Companies Share as a % of total companies, '25: 7.2%

Indicator 3: Data Suppliers' revenues

The aggregated value of all the data-related products and services generated by EU Data suppliers companies.

€ Bn | 2020: 90 | 2025: 101 | CAGR '25/'20: 2.4%

Indicator 4: Value of the Data Market

The marketplace where digital data is exchanged as "products" or "services" as a result of the elaboration of raw data.

€ Bn | 2020: 78 | 2025: 93 | CAGR '25/'20: 3.6%

Indicator 5: Value of the Data Economy

The indicator captures the potential gap between demand and supply of data skills in Europe.

€ Bn | 2020: 477 | 2025: 546 | CAGR '25/'20: 2.7% | Share of EU GDP, '25: 3.5%

Source: EDM Monitoring Tool, IDC 2019

Number of Data Suppliers

		2017	2018	Growth '18/'17
	Thousand	303,552	308,943	1.8%
	Thousand	36,906	37,344	1.2%
	Thousand	104,587	105,103	0.5%
	Thousand	271,700	272,200	0.2%

Value of the Data Economy

		2017	2018	Growth '18/'17	Share of GDP
	Direct Impacts € Million	146,966	162,157	10.3%	2017 1.% 2018 1.2%
	Backward Indirect Direct Impacts € Million	7,860	8,454	7.6%	Growth '18/'17 13.4%
	Direct Impacts € Million	6,996	7,593	8.5%	2017 0.2% 2018 0.2%
	Backward Indirect Direct Impacts € Million	335	352	5%	Growth '18/'17 27.4%

Value of the Data Market

		2017	2018	Growth '18/'17
	€ Million	146,970	162,154	10.3%
	€ Million	6,998	7,592	8.5%
	€ Million	26,720	29,297	9.6%
	€ Million	65,286	71,593	9.7%

		2017	2018	Growth '18/'17	Share of GDP
	Direct Impacts € Million	27,296	29,283	7.3%	2017 1% 2018 1.1%
	Backward Indirect Direct Impacts € Million	1,230	1,317	7%	Growth '18/'17 9.8%
	Direct Impacts € Million	65,286	71,593	9.7%	2017 0.5% 2018 0.5%
	Backward Indirect Direct Impacts € Million	3,316	3,972	19.8%	Growth '18/'17 7.9%

Source: EDM Monitoring Tool, IDC 2019

European Commission

IDC ANALYZE THE FUTURE

theLisboncouncil think tank for the 21st century

1. INTRODUCTION

The European Data Market Study (SMART 2013/0063) was launched by the European Commission in 2013 to measure the progress, size and trends of the European Data Economy with the objective of supporting the Data Value Chain policy of the European Commission. The study designed, developed and implemented a European Data Market Monitoring Tool providing facts and figures on the size and trends of the EU Data Market and Data Economy in the form of a series of quantitative indicators. The study also covered quali-quantitative aspects of the European Data Economy in the form of quantified stories investigating elements of the Data Market that were not captured by the Monitoring Tool. Finally, the European Data Market Study included a data landscaping tool offering a continuously updated picture of data companies in Europe and comprehended a series of webinars to disseminate the research results.

To continue gathering reliable and fact-based evidence on the EU Data Economy and measure the progress of the data-driven economy policies within the general framework of the Digital Single Market Strategy, the European Commission commissioned an update of the European Data Market (EDM) Study. The present document constitutes the **Second Interim Study Report (D2.6)** of the Update of the European Data Market Study (SMART 2016/0063), entrusted in 2016 to IDC and the Lisbon Council. As a follow-up to the First Interim Report (D2.3), this report brings together the research results and the activities carried out by the contractors under:

- **The Second Report on Facts & Figures (D2.4)** extending the measurement of the European Data Market Monitoring Tool by presenting data for the years 2016-2018 and forecasts to the year 2025 under three alternative scenarios;
- **The Second Report on Policy Conclusions (D2.5)** measuring the progress of European policies towards the objective of maximising the growth of the Data Economy as measured by the European Data Market Monitoring Tool;
- The key messages from the **quantified stories (D3.3, D3.4 and D3.5)** produced by the study team and focusing on the operational, organizational and/or economic benefits generated by the use of data-driven technologies with a special focus on Big Data and Artificial Intelligence (AI) solutions.
- **The Second Data Landscape Report (January 2019 Review – D4.2)** providing an overview of the EU Data Landscape and offering and up-to-date zoom into the database of data market companies in Europe.

1.1 Objectives

As for the previous study, the Update of the European Data Market Study (SMART 2016/0063) pursues three main objectives closely interrelated, which together allow to develop a complete and coherent picture of the European Data Market and Data Economy. They are as follows:

- Measuring the EDM indicators, providing facts and figures on all the key features of the European Data Market and Economy, regularly updated during the life of the project, building on the taxonomy and methodology approach previously developed and successfully implemented;
- Analysing relevant issues for the development of the data ecosystem, providing Data Market stories based on factual evidence, case studies and complementary data to the EDM indicators, following on the 12 stories already published by the previous study;

- Mapping and visualising the stakeholders populating the EU Data Market, building on the stakeholders' landscape and community developed in the previous study, and leveraging the visibility achieved by the website www.datalandscape.eu.

1.2 Methodological Approach

The Indicators

As outlined in the Second report on Facts & Figures (D2.4), the measurement of each of the indicators captured by the EDM Monitoring Tool is based on a sophisticated methodology combining data collection, data modelling, and desk research. Some initial assumptions are built on the original surveys completed when developing the initial European Data Market study (SMART2013/0063), and further supported by ongoing annual surveys among six Member States, namely the Big Six (France, Germany, Italy, Poland, Spain and the U.K.). While the initial surveys targeted potential data companies in two industries, namely ICT and Professional Services, and data users in 11 industries, the annual update surveys target all business sectors, and company sizes greater than 10 employees. The surveys are balanced to represent the mix of industries and size bands for companies in the European Union. The models used to represent expected market and company behaviour take inputs from macroeconomic indicators such as GDP and GDP growth, ICT spending, and employment. The main data sources used to compile the indicators are outlined in the table below.

Table 1: Main Data Sources by Indicator

Data Source	Updated	Used in
Eurostat Business Demographic Statistics	Dec 2018	Data professionals Data companies Data users
Eurostat annual structural business statistics	Dec 2018	Data professionals Data companies Data users
Eurostat chain linked Volumes (GDP)	Jan 2019	Data Market Data Revenues
IDC Core IT Spending guide 1Half 2018	Jun 2018	Data Market Data Revenues
IDC Worldwide Black Book[1] (standard edition)	Dec 2018	Data Market Data professionals Data Companies Data Users Data Revenues
IDC European Vertical Markets survey (2018)	Sep 2018	Data Market
IMF World Economic Outlook	Oct 2018	Data Market Data Revenues Data Economy
Consensus Forecasts – Consensus economics	Nov 2018	Data Market Data Revenues Data Economy
IT Big Data and Analytics spending Guide 2Half 2018	Jul 2018	Data Market
ILOSTAT statistics and databases	Jan 2019	Data Professionals

[1] IDC Black Books' series are the industry-standard study on the state of ICT spending in every region around the world. IDC's Black Books present a quarterly analysis of the size and growth of the worldwide ICT industry in 54 countries. https://www.idc.com/promo/customerinsights/blackbooks

Additional relevant sources leveraged for the measurement of the indicators were IDC Vertical Markets end user surveys and IDC Worldwide Black Book, whose results were used to confirm and adjust estimates, when necessary, of the number of companies that were data users and data suppliers.

The updated numbers of data users and data supplier companies were subsequently used to determine the updated results for the data companies' revenues and were further combined with above mentioned sources to measure the indicators for Data Professionals, Data Professionals' Skills Gap for the year 2017, 2018 and for the three 2025 scenarios.

The Second Report on Policy Conclusions

As a follow-up to the First Report on Policy Conclusions (D2.3), additional desk research and literature review were conducted to produce the Second Report on Policy Conclusions (D2.5) accompanying the quantitative results of the Second Report on Facts & Figures (D2.4). To better investigate the role of policies in shaping the current and future development of the European Data Economy, the study team leveraged a mix of IDC research and other sources. A select list of these sources is offered in Table 2 below:

Table 2: Main Sources

Document	Year	Author(s)
Digitising European Industry , COM(2016) 180 final	2016	European Commission
Artificial Intelligence: the next digital frontier?	2017	Mc Kinsey Global Institute
Developing Supercomputers in Europe	2017	European Parliamentary Research Service
Artificial Intelligence for Europe, COM (2018) 237 final	2018	European Commission
Coordinated Plan on Artificial Intelligence, COM(2018) 795 final	2018	European Commission
Towards a common European data space, COM(2018) 232 final	2018	European Commission
The global policy response to AI	2018	FTI Consulting
Artificial Intelligence for the Real World	2018	Davenport T. H., Ronanki R.
AI In Europe: Key Findings of IDC's 2018 AI User Survey (Doc #EMEA44220518)	2018	IDC
The Global Competitiveness Report	2018	Schwab K., World Economic Forum
How do you build a successful Data-as-a-service business?	2018	Dan Vesset, IDC, Group VP Analytics
Western Europe risks losing the Technology Race	2018	Marc Dowd and Mark Yates, IDC
Artificial Intelligence: A European Perspective	2018	European Commission's Joint Research Centre (JRC)
Digital Economy and Society Index (DESI) 1 2018 - Country Report United Kingdom	2018	European Commission
Future of Work, Future of Society	2018	EGE European Group on Ethics in Science and New Technologies
IDC Worldwide Semiannual Artificial Intelligence Systems Spending Guide	2019	IDC

Document	Year	Author(s)
Ethics Guidelines for Trustworthy AI	2019	High Level Experts Group on AI
Regulation on the Free Flow on Non-Personal Data	2019	European Commission
Data Protection and Brexit	2019	ICO
EU-Japan Economic Partnership Agreement	2019	EU - Japan

The Quantified Stories

The quantified stories were the result of a mixed effort entailing both secondary and primary research activities. Extensive secondary research on available public sources, specialised press and academic literature was undertaken to obtain an actionable and up-to-date understanding of private and scientific data for public interest and innovation together with a comprehensive picture of the phenomenon in Europe and worldwide.

In parallel, primary research was conducted to collect empirical evidence and validate the information obtained through the main desk research activities. Among the organisations interviewed, the following featured a prominent role:

For the story on data monetization:

- DAWEX, Information Technology / Start Up (France)
- Pirelli, Manufacturing sector (Italy)
- TfS (Transport for London), Local government / Transport (U.K)
- BT (British Telecom), Telecommunications (U.K)
- Bosch, Automotive parts, power tools, security systems, home appliance, engineering, electronics, motorized bicycle motors (Germany)

For the story on how Big Data is driving Artificial Intelligence across different sectors, we have further carried out in-depth interviews with representatives of DataIku - a leading AI and Machine Learning technology provider representing a central hub for analytics and machine learning model deployment and management. These secondary and primary research efforts led to the realization of five distinct case-studies in different sectors and across different AI use cases. In particular:

- Coyote System SAS – a France-based real-time road information provider
- Voyage Privé SAS – an online travel agency
- DAZN – a live-streaming service specialized in sporting events
- AXA –a French multinational insurance firm –
- Airbus – the aerospace company.

For the story on how Artificial Intelligence is implemented across Europe's utilities, we have carried out in-depth interviews with representatives of Atos – a European multinational information technology (IT) service and consulting company headquartered in France. These secondary and primary research efforts led to the realization of five distinct case-studies in different sectors and across different AI use cases. In particular:

- Enel Green Power (EGP), the renewables divisions of Italy-based energy utility Enel on solar plant maintenance;
- E.ON, one of Germany's largest electric utility companies based North Rhine-Westphalia, on predictive maintenance and energy efficiency as-a-service;

- Innogy, the renewables, network, and retail businesses of German electric utility RWE, on intelligent customer inquiry;
- Fortum, a Finland-based international energy utility company, on intelligent heating management.

The Second EU Data Landscape Report

The Second EU Data Landscape Report (D4.2) provides a detailed overview of the updated EU Data Landscape database and a zoom into the stakeholders and their positioning in the data economy environment. Following the January 2018 update reported in the First EU Data Landscape report (D4.1), the EU Data Landscape database has been revised to capture the current trends and include the data collected in the period from January 2018 until January 2019.

The dataset has been significantly extended through desk research as well as through input received from stakeholders in the data economy, for instance, via the www.datalandscape.eu website. The report relies on the crowdsourcing of knowledge through an open process, where stakeholders can directly suggest the companies to be included in the database. The mapping exercise sought to achieve a balanced and comprehensive coverage of the different geographies, different typologies of companies (SMEs, large companies, research institutions etc.) and the different data sectors. In terms of geographical coverage, the mapping of the Data Landscape focuses on the EU 28. However, companies from other European countries as well countries outside Europe are also depicted in the database. The reviewing procedure consisted of the following steps:

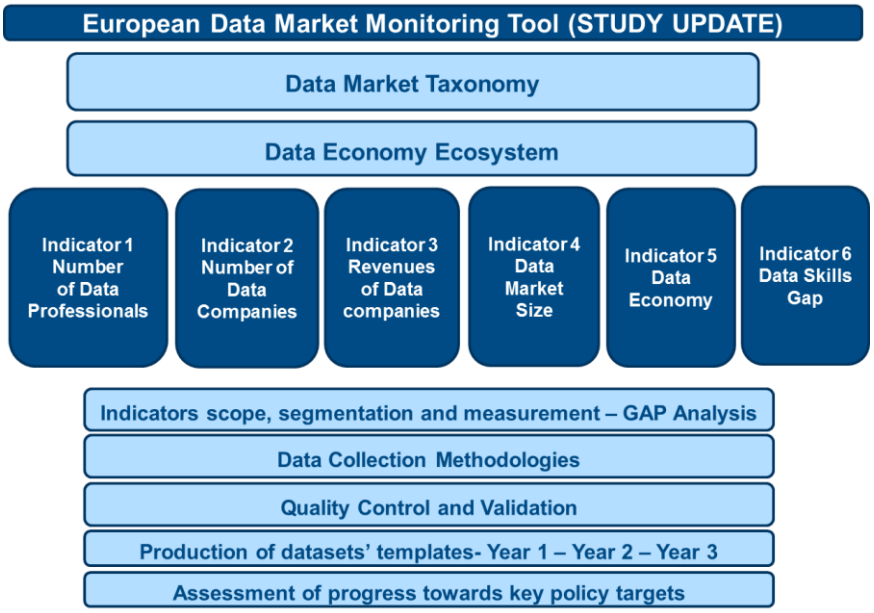
- Control of the existing dataset;
- Extension of the dataset, taking into account coverage goals;
- Review of Key Data Landscape companies and identification of new.

Finally, among the 1459 companies included in the dataset, 259 have been identified as Key Data Landscape companies in line with a set of criteria adopted.

1.3 The European Data Market Monitoring Tool

In line with the results presented in the original European Data Market study (SMART 2013/0063) in February 2017, the First Report on Facts & Figures (D2.1) of February 2018 and the Second Report on Facts & Figures (D2.4) of March 2019, the indicators presented in this report are organised around a modular and flexible structure – the European Data Market Monitoring Tool. The updated European Data Market Monitoring Tool designed by IDC is shown in the Figure below and its main components are further described in the following sections.

Figure 5: The Updated EDM Monitoring Tool



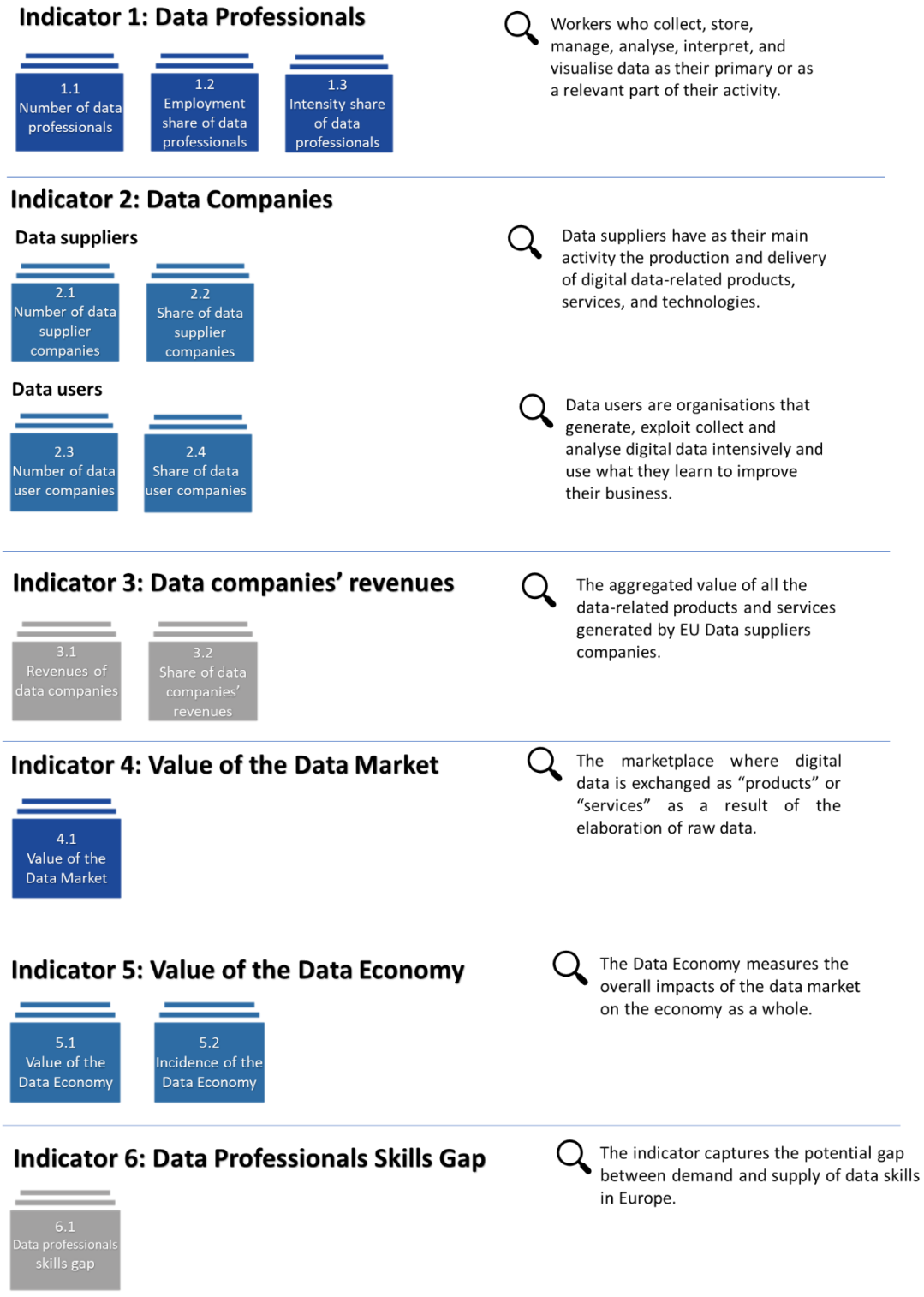
1.4 The Structure of this Report

The present report is built along the following sections:

- The first section – corresponding to Chapter 2 – summarises the results of the Second Report on Facts & Figures (D2.4) that was delivered and approved by the European Commission in March 2019.
- The second section – corresponding to Chapter 3 – provides additional qualitative and quantitative aspects on the European Data Market as obtained by the quantified stories (D3.3, D3.4 and D3.5) produced by the study team between August 2018 and May 2019.
- The third section – corresponding to Chapter 4 – presents an updated overview of the data landscape and interactive Data Market Monitoring Tool based on the January 2019 update reported in the Second Data Landscape Report (D4.2).
- The fourth section – corresponding to Chapter 5 – focuses on the policy conclusions delivered in the Second Report on Policy Conclusions (D2.5) of May 2019 and approved in June 2019.
- The final section provides for a set of concluding remarks drawing from all the different components (and deliverables) of the Update of the European Data Market study.

2. QUANTIFYING THE DATA MARKET – KEY FACTS & FIGURES

The key facts & figures stemming from the second round of measurement of the Update European Data Market Study (SMART 2016/0063) as reported in the Second Report of Facts & Figures were obtained through the measurement of the following set of selected indicators:



Each indicator was measured at the level of the total EU28 and total EU27 (excluding the U.K.) for all 28 EU Member States, when available and applicable; industry-specific and company-size views were also offered with indicators provided by industry sector and company size bands, when possible. As in the European Data Market Study (SMART 2013/0063), a select number of indicators has been developed and updated for three non-European countries, namely Brazil, Japan and the United States.

The six key indicators measured by the EDM can be seen holistically along four main dimensions:

- The Workforce and Skills dimension - including the measurement of data professionals and their potential skill gap.
- The Supply and Demand dimension - incorporating the measurement of data supplier and data user companies and the revenues generated by data supplier companies.
- The Business and Economy dimension - comprehending the size of the Data Market and the value of the Data Economy.
- The International context dimension - including a select number of indicators for Brazil, Japan and the U.S.

Figure 6: The four Dimensions of the Data Market's Key Facts & Figures

Source: The European Data Market Monitoring Tool, IDC, 2019

2.1 Three future Development Paths: The Data Market at 2025

The key facts & figures obtained through the measurement of the above-listed indicators are presented for the years 2017 and 2018 as well as for the year 2025 according to three potential future scenarios of the European Data Market and Economy, driven by different macroeconomic and framework conditions. The scenarios at 2025 continue to take as a reference point the initial scenarios developed for the year 2020. While the 2020 scenarios were mainly differentiated by economic drivers (different demand-supply dynamics), the 2025 scenarios continue to be shaped by a combination of economic and social drivers, focused on the interaction of two main focal issues (or evolution paths):

- the high or low pace of diffusion of data-driven innovation, driven by demand-supply dynamics, and its impact on economic growth;
- the high or low concentration of power in the access, control and exploitation of data assets, that is the social model of data governance. At one extreme, we foresee a society where a few actors, such as leading online platforms, governments, large businesses, dominate the

main data assets and therefore are able to capture a disproportionately high share of data innovation benefits, increasing social inequality (highly centralized model). The polar opposite of this scenario would be a society characterised by an open, transparent and participatory approach to data governance, where both citizens and organisations are able to control and extract value from their data. This would result in a wider social distribution of data innovation benefits, decreasing social inequality.

This analysis highlights the critical turning points to be faced in the next years by governments, businesses and social actors in the development of the European Data Economy. The combination of alternative social and economic trends results in the following scenarios:

- The **Baseline scenario** is characterised by a healthy growth of data innovation, a moderate concentration of power by dominant data owners with a data governance model protecting personal data rights, and an uneven but rather wide distribution of data innovation benefits in the society;
- The **High Growth scenario** is characterised by a high level of data innovation, low data power concentration, an open and transparent data governance model with high data sharing, and a wide distribution of the benefits of data innovation in the society;
- The **Challenge scenario** is characterised by a low level of data innovation, a moderate level of data power concentration due to digital markets fragmentation, and an uneven distribution of data innovation benefits in the society.

These scenarios underline that there are relevant choices to be made in the next years about the social and economic governance model of the Data Market, in order to maximise the chances of harnessing the power of data for economic growth together with an open, transparent and shared model of data governance and control.

2.2 The Workforce Dimension: Data Professionals and Data Skills Gap

Measuring the Data Professionals

Data professionals[2] *are workers who collect, store, manage, and/or analyse, interpret, and visualise data as their primary or as a relevant part of their activity. Data professionals must be proficient with the use of structured and unstructured data, should be able to work with a huge amount of data and be familiar with emerging database technologies.*

Data Professionals in 2017 and 2018

According to the second round of measurements of the EDM, data professionals are estimated at a total of 5.7 million in the EU27 and at 7.2 million in the EU28 in 2018, thus marking a significant increase in 2018 over the previous year (8.6% and 8.4% year-on-year respectively). When compared to the year 2018, year 2020 would register a Compound Annual Growth Rate (CAGR) of 7.2% and 6.9% at the level of EU27 and EU28 respectively. More interestingly, the employment share and the

[2] The previous European Data Market Study (SMART 2013/0063) included an indicator measuring "Data Workers", which was based on a similar, but slightly more restrictive definition. In this updated study we have decided to measure "Data Professionals", that is workers with a wider range of data-related roles. Indeed, data professionals are not only data technicians, but also users who, based on sophisticated tools, take decisions about their business or activities after having analysed and interpreted available data.

intensity share components of the data professionals' indicator are also expected to significantly improve in 2018 and 2020 if compared to our estimates in 2016 (now estimated at 3% and 3.4% in 2017 and 2020 in the EU27 and 3.3% and 3.7% for the same years in the EU28). As underlined in the First Report on Facts & Figures, this increase confirms the positive evolution of the workforce involved in data-related professions over the period under consideration.

Table 3: Data Professionals, 2016-2017-2018-2020 and Growth Rates

Indicator 1 — Data Professionals 2016-2017-2018-2020 and Growth Rates									
N.	Region	Name	Description	2016	2017	2018	2020	Growth 2018/2017	CAGR 2020/2018
1.1	EU27	Number of data professionals	Total number of data professionals in EU (000s)	4,875	5,260	5,713	6,565	8.6%	7.2%
1.1	EU28	Number of data professionals	Total number of data professionals in EU (000s)	6,187	6,666	7,226	8,251	8.4%	6.9%
1.2	EU27	Employment share of data professionals	Share of data professionals on total employment in EU (%)	2.8%	3.0%	3.2%	3.4%	5.6%	6.8%
1.2	EU28	Employment share of data professionals	Share of data professionals on total employment in EU (%)	3.1%	3.3%	3.4%	3.7%	5.5%	6.6%
1.3	EU27	Intensity share of data professionals	Average number of data professionals per user company (units)	9.6	10.2	11.0	12.2	8.6%	9.5%
1.3	EU28	Intensity share of data professionals	Average number of data professionals per user company (units)	9.2	9.6	10.4	11.4	8.4%	8.9%

Source: European Data Market Monitoring Tool, IDC 2019

Data Professionals at 2025

A steady progression of the number of data professionals continues to emerge from our latest estimates. The number of data professionals in both the EU27 and EU28 is forecast to grow significantly under all the three scenarios at 2025 as the use of data-driven innovation is expected to grow unabated even under the less economically favourable scenario. In particular, under the Baseline scenario, data professionals are expected to amount to 9.3 million in the EU27 and 11.3 million in the EU28 at 2025, thus representing a solid growth rate between 7.3% and 6.6% over the 2018-2025 period. In the Challenge and High Growth scenarios, data professionals would be more than 8.6 million and 10.8 million in the EU27 and 10.3 million and 13.1 million in the EU28 respectively. Under all three scenarios, the CAGR over the period 2018-2025 is consistent, although higher than the CAGR featured by the Data Market growth, thus confirming again the close relationship between the two variables.

Table 4: Data Professionals Forecast 2025 - Total Number in the EU27 and EU28 and Growth Rates. Challenge, Baseline and High Growth Scenarios (Units, '000; %)

Indicator 1 — Data Professionals – Forecast 2025									
N.	Region	Name	Description	2025 Challenge	2025 Baseline	2025 High Growth	CAGR Challenge scenario	CAGR Baseline scenario	CAGR High Growth scenario
1.1	EU27	Number of data professionals	Total number of data professionals in EU (000s)	8,597	9,320	10,888	5.5%	7.3%	10.6%
1.1	EU28	Number of data professionals	Total number of data professionals in EU (000s)	10,346	11,339	13,164	4.6%	6.6%	9.8%

Source: European Data Market Monitoring Tool, IDC 2019

Measuring the Data Professionals Skills Gap

*The **Data Professionals Skills Gap** indicator captures the potential gap between demand and supply of data skills in Europe, since the lack of skills may become a barrier to the development of the data industry and the rapid adoption of data-driven innovation. It is based on a model balancing the main sources of data skills (from the education system and re-training and other carriers) with the estimated demand (by all data companies).*

This indicator continues to signal an imbalance between demand and supply of data skills in Europe since the first measurement for the year 2014. In the year 2018 the strong increase of demand of data professionals continued (+7%), the estimated gap grew by 10% reaching approximately 571,000 unfilled positions in the EU28 corresponding to 7.2% of total demand. By 2020 we expect the gap to expand to 641,000 unfilled positions in EU28, corresponding to 7.8% of total demand. In any given moment in the labour market there is a physiological number of vacancies, as well as of people looking for work: a vacancies ratio around 5% of demand or less is considered manageable. From this point of view the data skills gap estimated for 2018 shows stress in the market, which is expected to continue to 2020 and beyond under the 3 scenarios.

The three forecast scenarios all predict an increasing data skills gap, which is worrying based on the criteria defined above. The forecast gap ranges from 9% of demand in the Challenge one, to 10% in the Baseline, to 14% in the High Growth (EU27), as shown in the following summary table. The slightly lower dynamic of data skills demand in the U.K. (due to slightly slower forecast of the data market growth) means that the gap share of demand is smaller in EU28 than in EU27.

The absolute size of the data skills gap is relevant, potentially reaching 925,000 unfilled positions in 2025 in the EU 27 Baseline scenario, but up to over 1.5 million in the EU 27 High Growth scenario. In the Challenge scenario the data skills gap is forecast at 775,000 unfilled positions in 2025. This underlines the need for policy action to prevent and minimize the unbalance between data skills demand and supply in the next years.

Table 5: Indicator 6 - Data Professionals Skills Gap in the EU, 2017-2018-2020 and 2025 - Three scenarios

Indicator 6 — Data Professionals Skills gap												
N.	Name	Description		Actual		Baseline Scenario	Baseline Scenario		Challenge Scenario		High Growth Scenario	
				2017	2018	2020	2025	2018 - 2025 CAGR	2025	2018 - 2025 CAGR	2025	2018 - 2025 CAGR
6.1	Data professionals' skills gap	Demand / Supply Gap N, 000s	EU27	395	496	603	925	7.0%	775	6.6%	1,551	17.7%
			EU28	483	571	641	1039	9.2%	829	5.5%	1,693	16.8%
6.2		Demand / Supply Gap % of demand	EU27	6.7%	7.9%	9.1%	9.9%		9.0%		14.2%	
			EU28	6.5%	7.2%	7.8%	9.2%		8.0%		12.9%	

Source: European Data Market Monitoring Tool, IDC 2019

2.3 The Supply - Demand Dimension: The Data Companies

Measuring the Data Companies

Data companies *are organisations that are directly involved in the production, delivery and/or usage of data in the form of digital products, services and technologies. They can be both data suppliers' and data users' organisations:*

- ***Data suppliers*** *have as their main activity the production and delivery of digital data-related products, services, and technologies. They represent the supply side of the Data Market.*
- ***Data users*** *are organisations that generate, exploit, collect and analyse digital data intensively and use what they learn to improve their business. They represent the demand side of the Data Market.*

Data Companies in 2017 and 2018

The number of data suppliers continue to grow at a faster pace than the numbers of data users in the longer term – out to 2025. Data suppliers are estimated at almost 145,100 in the EU27 and at more than 283,000 units in the EU28 for 2018, thus exhibiting a year-on-year growth of 4.1% and 4.2% respectively. Data users, instead, are projected to grow at 3.4% in 2018, amounting to more than 534,000 in the EU27 and to 715,000 units in the EU28. If compared to the measurements carried out by the European Data Market Monitoring Tool over the period 2013-2015, these latest estimates show a picture of consolidation of data companies in the EU, following increasing growth rates increasing over the prior four years.

This consolidation is reflected in stable share of data companies over the total number of companies in Europe. The share of data suppliers on total companies in the ICT and Professional services industries is estimated at 11.6% in the EU27 and 15.3% in the EU28, a slight improvement with respect to 2017. The data users' penetration rates (i.e. the share of data users on total companies in the EU) are also stable with a modest percentage increase in 2018 in both the EU27 and the EU28.

Table 6: Data Companies, 2016-2017-2018-2020 and Growth Rates

N.	Name	Description	Market	2016	2017	2018	2020	Growth 2018/2017
2.1	Number of data suppliers	Total number of data suppliers measured as legal entities based in the EU (000s)	EU27	134,300	139,450	145,100	150,800	4.1%
2.1	Number of data suppliers	Total number of data suppliers measured as legal entities based in the EU (000s)	EU28	261,450	271,700	283,100	294,350	4.2%
2.2	Share of data suppliers	% share of data companies on total companies in the ICT and Professional services industries	EU27	10.9%	11.3%	11.6%	12.1%	2.6%
2.2	Share of data suppliers	% share of data companies on total companies in the ICT and Professional services industries	EU28	14.2%	14.8%	15.3%	16.1%	3.3%
2.3	Number of data users	Total number of data users in the EU, measured as legal entities based in one EU country	EU27	505,950	517,100	534,450	538,600	3.4%
2.3	Number of data users	Total number of data users in the EU, measured as legal entities based in one EU country	EU28	676,150	691,500	715,000	721,300	3.4%
2.4	Share of data users	% share of data users on total companies in the EU industry	EU27	5.7%	5.8%	6.0%	6.0%	3.4%
2.4	Share of data users	% share of data users on total companies in the EU industry	EU28	6.5%	6.6%	6.8%	6.9%	3%

Source: European Data Market Monitoring Tool, IDC 2019

Data Suppliers Forecasts at 2025

According to our latest forecasts at 2025 as far as data suppliers are concerned, the outlook for the number of data suppliers is on continued increase beyond 2020. A slight reduction vis-à-vis our previous estimates is to be observed – primarily because the market is larger and so further growth becomes more difficult. Despite this, there is much higher growth for the larger data supplier companies because investment as a data supplier requires resources not as readily available to smaller companies.

Larger companies can afford specialised staff and departments whose sole purpose is to address the Data Market, while in smaller companies the development role often falls to individuals who have other responsibilities.

Table 7: Data Suppliers Forecast 2025 by Member State - Three Scenarios (Units; '000); CAGR 2025-2020 (%)

Indicator 2 – Data Suppliers – Forecast 2025						
Region	2025 Challenge	2025 Baseline	2025 High Growth	CAGR 2025/2020 Challenge Scenario (%)	CAGR 2025/2020 Baseline Scenario (%)	CAGR 2025/2020 High Growth Scenario (%)
EU27	163,350	174,500	194,450	1.6%	3.0%	5.2%
EU28	320,700	338,700	389,100	1.7%	2.8%	5.7%

Source: European Data Market Monitoring Tool, IDC 2019

Data Users Forecasts at 2025

Long term growth in the number of data user companies is highest in the data intense industries such as Professional services and ICT, and lowest in Mining and Construction. In line with the results from the First Report on Facts & Figures, the largest companies show the highest growth in adoption as the Data Economy will be crucial to their success and competitive advantage – without a data-oriented approach to business and business decisions these companies will not see the opportunities their competitors see and so will not grow at the same rate. However, these larger companies are a small share of the overall number of companies, so, although the number will grow at a compound rate of 27% to 2025 compared with 1.0% for those in the smaller size band, they do not add significantly to the total number of data companies.

Table 8: Data Users Forecast 2025 by Member State - Three Scenarios (Units; '000); CAGR 2025-2020 (%)

Indicator 2 – Data Users – Forecast 2025						
Region	2025 Challenge	2025 Baseline	2025 High Growth	CAGR 2025/2020 Challenge Scenario (%)	CAGR 2025/2020 Baseline Scenario (%)	CAGR 2025/2020 High Growth Scenario (%)
EU27	558,550	580,950	624,400	0.7%	1.5%	3.0%
EU28	750,800	778,600	844,250	0.8%	1.5%	3.2%

Source: European Data Market Monitoring Tool, IDC 2019

Measuring Data Companies' Revenues

Data companies' revenues *correspond to the aggregated value of all the data-related products and services generated by Europe-based data suppliers, including exports outside the EU.*

Data Companies' Revenues in 2017 and 2018

Revenues generated by data suppliers have registered a constant increase since 2013 according to our initial measurements and the Monitoring Tool. In 2018, in particular, revenues have increased by more than 12% to reach nearly 59 Billion Euro in the EU27 and 77 Billion Euro in the EU28. The share of the data suppliers' revenues on the total companies' revenues in the ICT and Professional services sectors rose to 3.4% in the EU27 and 3.5% the EU28 in 2018.

Table 9: Data Companies' Revenues and Growth, 2016-2017-2018-2020 (€, Million; %)

Indicator 3 — Data Companies' Revenues and Growth									
N.	Region	Name	Description	2016	2017	2018	2020	Growth 2018/2017	CAGR 2020/2018
3.1	EU27	Total revenues of data companies in the EU	Total revenues of the Data Suppliers calculated by Indicator 2	47,178	52,479	58,978	69,450	12.4%	8.8%
3.1	EU28	Total revenues of data companies in the EU	Total revenues of the Data Suppliers calculated by Indicator 2	61,781	68,846	77,082	89,263	12.0%	7.9%
3.2	EU27	Share of data companies' revenues	Ratio between Data Suppliers' revenues and total companies' revenues in sectors J and M	3.0%	3.2%	3.4%	NA	7.0%	NA
3.2	EU28	Share of data companies' revenues	Ratio between Data Suppliers' revenues and total companies' revenues in sectors J and M	3.1%	3.3%	3.5%	NA	8.2%	NA

Source: European Data Market Monitoring Tool, IDC 2019

Data Companies' Revenues Forecasts at 2025

Data companies' revenues are expected to follow the Data Market, as imports and exports of data tools and services tend to follow each other. Forecasting data companies' revenues shows an expected annual growth rate out to 2025 of 7.1% - easily outpacing the growth of the total ICT market over the same period (expected to be 1.5% from 2020 to 2025 Baseline). The smaller Member States show the highest long-term growth as they have a smaller base from which to grow,

but the larger Member States will make the biggest overall contribution to the Data Economy out to 2025.

Table 10: Data Companies' Revenues Forecast 2025 by Member State - Three Scenarios (€, Million; %)

Indicator 3 — Data Companies' Revenues - Forecast 2025						
	2025 Challenge	2025 Baseline	2025 High Growth	CAGR 2025/2020 Challenge Scenario (%)	CAGR 2025/2020 Baseline Scenario (%)	CAGR 2025/2020 High Growth Scenario (%)
Total EU27	77,622	97,108	133,755	2.1%	6.8%	13.9%
Total EU28	101,117	126,342	178,879	2.4%	7.1%	14.8%

Source: European Data Market Monitoring Tool, IDC 2019

2.4 The Business and Economic Dimension: The Data Market and the Data Economy

Measuring the Data Market

*The **Data Market** is the marketplace where digital data is exchanged as "products" or "services" as a result of the elaboration of raw data.*

The Data Market in 2017 and 2018

The value of the Data Market in 2018 for both the EU27 and the EU28 is showing a growth rate well above the one exhibited by the Total IT spending, at 9.7% year-on-year and is expected to surpass the threshold of 60 Billion Euro in 2020 in the EU27 – a constant and significant progression if we consider that the total amount of the Data Market in the EU27 was estimated at 42.6 Billion Euro in 2015 in our previous study, and that our current estimates measure the Data Market at 50.6 Billion Euro in 2017.

Table 11: Data Market Value and Growth, 2016-2017-2018-2020 (€, Million; %)

Indicator 4 — Value and Growth of the Data Market									
N.	Market	Name	Description	2016	2017	2018	2020	Growth 2018/2017	CAGR 2025/2018
4.1	EU27	Value of the Data Market	Estimate of the overall value of the Data Market	46,183	50,604	55,511	60,340	9.7%	4.3%
4.1	EU28	Value of the Data Market	Estimate of the overall value of the Data Market	59,496	65,286	71,593	77,769	9.7%	4.2%

Source: European Data Market Monitoring Tool, IDC 2019

The Data Market Forecasts at 2025

Our latest estimates of the Data Market value in 2025 under the High Growth Scenario continue to showcase a buoyant growth of almost 200% with respect to the estimates for the year 2018 in both the EU27 and the EU28. This will correspond to a CAGR for the period 2020-2025 of 12.1% and 12.7% in the EU27 and the EU28 respectively. According to our new 2025 Baseline scenario, the Data Market will amount to more than 82 billion Euro in the EU27, against 60.3 billion Euro in 2018 (a 6.5% CAGR 2020-2025), while under the Challenge scenario the Data Market will still represent 72.1

billion Euro, growing at a compound annual growth rate of 3.6% from 2020. The Data Market growth will therefore continue unabated in 2025, confirming the trend set out in 2013-2014 while elaborating our initial results of the European Data Market Study (SMART 2013/0063). However, these forecasts for 2025 are slightly lower than the previous forecast, accounted for partly by increased pessimism for a general economic slowdown as well for the uncertainty surrounding Brexit.

Table 12: Data Market Forecast 2025 by Member State - Three Scenarios (€, Million; %)

Indicator 4 — Data Market - Forecast 2025						
	2025 Challenge	2025 Baseline	2025 High Growth	CAGR 2025/2020 Challenge Scenario (%)	CAGR 2025/2020 Baseline Scenario (%)	CAGR 2025/2020 High Growth Scenario (%)
Total EU27	72,099	82,779	106,969	3.6%	6.5%	12.1%
Total EU28	93,026	106,127	141,596	3.6%	6.4%	12.7%

Source: European Data Market Monitoring Tool, IDC 2019

Measuring the Data Economy

*The **Data Economy** measures the overall impacts of the Data Market on the economy as a whole. It involves the generation, collection, storage, processing, distribution, analysis elaboration, delivery, and exploitation of data enabled by digital technologies.*

The Data Economy includes the direct, indirect, and induced effects of the Data Market on the economy.

- *The direct impacts: are the initial and immediate effects generated by the data suppliers; they represent the activity potentially engendered by all businesses active in the data production. The quantitative direct impacts will then be measured as the revenues from data products and services sold, i.e. the value of the Data Market.*
- *The indirect impacts: are the economic activities generated along the company's supply chain by the data suppliers. There are two different types of indirect impacts: the backward indirect impacts and the forward indirect impacts.*
- *The induced impacts: include the economic activity generated in the whole economy as a secondary effect.*

The Data Economy in 2017 and 2018

The value of the Data Economy for the EU28 has been estimated to exceed the threshold of 300 Billion Euro in 2018, overall confirming the estimated growth of nearly 12% presented in the previous deliverable. The estimated share of overall impacts on GDP is 2.4% in 2017 and is expected to grow to 2.6% in 2018. Brexit uncertainties play an important role in affecting the results for EU28, which shows lower 2018 growth as well as lower 2017-2020 CAGR, now even more than during the previous year. Indeed, since no deal has been reached yet, only two months before the Brexit is expected to be completed.

Table 13: Data Economy Value and Growth, 2017-2018-2020 and Impacts on GDP 2017-2018 (€, Million; %)

Indicator 5 — Value and Growth of the Data Economy											
N.	Market	Name	Description	2016	2017	2018	2020	Growth 2018/ 2017	CAGR 2020/ 2018	Impacts on GDP 2017	Impacts on GDP 2018
5.1 5.2	EU27	Value of the Data Economy and Impacts on EU GDP	Value of direct, indirect and induced impacts on the EU economy and % of EU GDP	238,699	267,986	301,713	387,646	12.6%	13.3%	2.2%	2.4%
5.1 5.2	EU28	Value of the Data Economy and Impacts on EU GDP	Value of direct, indirect and induced impacts on the EU economy and % of EU GDP	299,989	336,602	376,925	477,297	12.0%	12.5%	2.4%	2.6%

Source: European Data Market Monitoring Tool, IDC 2019

The Data Economy Forecasts at 2025

When comparing the Baseline scenario at 2025 to the Data Economy estimated in 2018, we can see that the value will nearly double along the period, from around 377 Billion Euro in 2018 to nearly 680 Billion Euro in 2025 for EU28, with a CAGR 2018/2020 of 12.5% for EU28 and 13.3% per EU27, lowering down when considering the CAGR 2020/2025 (7.3% EU28 and 7.5% EU27). These results are indeed affected by the new macroeconomic forecasts, expecting economic growth to be slower from 2020, reflected in a lower 2020/2025 CAGR.

The CAGR 2020/2025 in the EU28 for the High Growth scenario is 17.2% for EU28 that will make the Data Economy for EU28 surpass the threshold of 1000 Billion Euro, and accounting for 6.3% of the GDP at 2025.

Similarly, the Challenge scenario will see a slowdown of the economic effects, from 2.8% CAGR 2020/2025, with the Data Economy being just above 500 Billion Euro. We consider this scenario as the most likely, considering the economic slowdown, with macroeconomic and political factors highly correlated and showing their effects after 2020.

Table 14: Data Economy Forecast in 2025 and Impacts on GDP according to the Three Scenarios (€, Million; %)

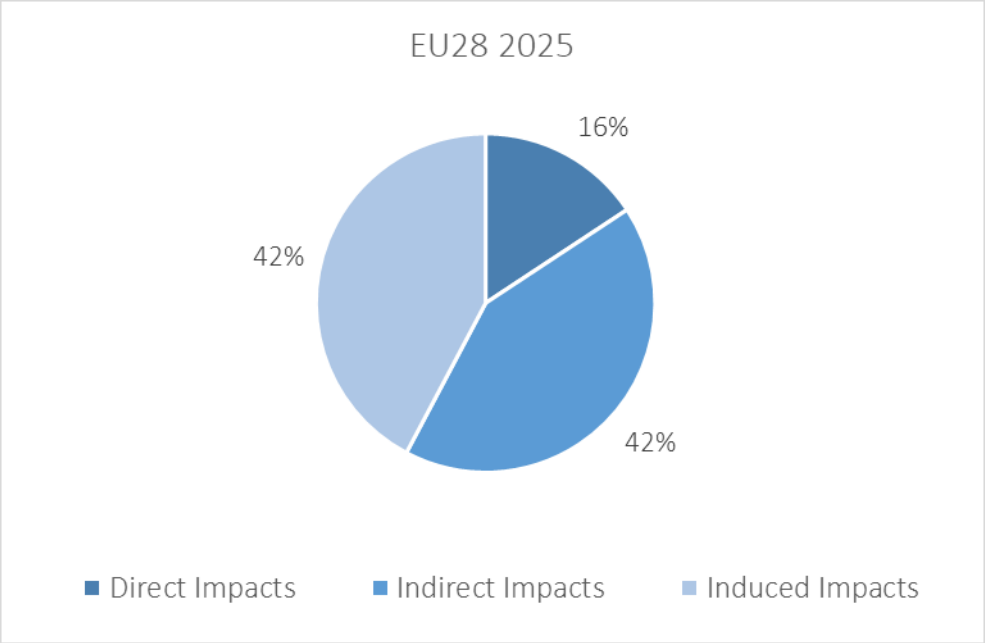
Indicator 5 – Data Economy - Forecast 2025								
N.	Name	Description	2025 Challenge Scenario	2025 Baseline Scenario	2025 High Growth Scenario	Impacts on GDP 2025 Challenge Scenario	Impacts on GDP 2025 Baseline Scenario	Impacts on GDP 2025 High Growth Scenario
5.1 5.2	Value of the Data Economy EU27	Value of direct, indirect and induced impacts on the economy	444,048	555,263	829,014	3.3%	4.0%	5.8%
5.1 5.2	Value of the Data Economy EU28	Value of direct, indirect and induced impacts on the economy	545,497	680,149	1,053,839	3.5%	4.2%	6.3%

Source: European Data Market Monitoring Tool, IDC 2019

As in the previous study, this report provides a detailed insight of the Data Economy by type of impact – direct, indirect and induced impacts. It is worth highlighting how the composition of impacts changes along time, from 2018 to 2025, in favour of induced impacts, thus revealing the effects of data access, data product and services exchange, and data value distribution in the economy.

Indeed, induced impacts in 2025 account for a share of 42% (both in EU27 and in EU28) from 33% in 2018. Indirect impacts in turn will lose around 10% of share, but still in 2025 accounting for a very high percentage (42%). With respect to 2018, in which the indirect impacts are the most relevant, forward impacts in particular, in 2025 induced impacts will increase, reaching a share similar to the one of the indirect impacts.

Figure 7: Data Economy by Type of Impact, EU28, 2025 (%)



Source: European Data Market Monitoring Tool, IDC January 2019

2.5 The International Dimension - The Data Economy Beyond the EU – US, Brazil and Japan

The U.S.

The number of data professionals, as well as their share on the country's total employment base, again marks the strongest growth in 2018 when compared with the other two non-EU countries assessed by the study, Brazil and Japan (4.6% and 2.9% year-on-year growth in the number of data professionals and employment share in 2018 over 2017 respectively). However, this growth has slowed when compared to 2017. The same applies for the data supplier companies' indicators, with the highest increase of data suppliers in 2017 (2.9% vs. 1.2% and 0.5% in Brazil and Japan respectively) and an even stronger rise in their associated revenues and in the value of the Data Market in 2018 over 2017 (10.3% vs. 8.5% and 9.7% in Brazil and Japan respectively). Accordingly, the U.S.' overall Data Economy (direct and backward indirect impacts only) increased by more than 10% in 2018 gaining a 13.4% year-on-year increase in incidence on GDP and now representing more than 0.8% of the country's GDP.

Table 15: USA Indicators - Overview 2016-2018

USA – Indicators' Overview						
N.	Name	Metrics	2016	2017	2018	Growth 2018/2017
1.1	Number of Data professionals	Total Number of Data professionals (Thousands)	12,732	13,857	14,497	4.62%
1.2	Data professionals' employment share	% of Data professionals on total employment	8.42%	9.04%	9.30%	2.90%
2.1	Number of Data Suppliers	Total number of data supplier companies (000s)	289,556	303,552	308,943	1.78%

USA – Indicators' Overview						
N.	Name	Metrics	2016	2017	2018	Growth 2018/2017
3.1	Revenues of Data Companies	Total revenues generated by companies specialized in the supply of data-related products and services (Million €)	129,173	146,970	162,154	10.33%
4.1	Value of the Data Market	Estimate of the overall a value of the Data Market (Million €)	129,173	146,970	162,154	10.33%
4.2	Value of the Data Economy (Only Direct and Backward Indirect impacts)	Direct Impacts (Million €)	108,521	146,966	162,157	10.34%
		Backward Indirect Impacts (Million €)	7,270	7,860	8,454	7.56%
4.3	Incidence of the Data Economy on GDP (Only direct and backward indirect impacts)	Ratio between value of the Data Economy and GDP (%)	0.78%	1.03%	1.17%	13.42%

Source: European Data Market Monitoring Tool, IDC 2019

Brazil

Brazil shows some economic recovery in 2018, with increased confidence in the economy, and a growth in IT spending in 2018 ahead of Japan, but still a little behind the U.S. Most of the indicators pertaining to the Data Market and the Data Economy grew in 2018 and are expected to continue on this journey in the years to come. While data professionals and their associated share on total employment declined between 0.7% and 1.4%, the number of data supplier companies, their related revenues and the accompanying Data Market grew at a rate comparable with the U.S. – at 1.2% and 8.5%. As a result, the incidence of the economy on Brazil's GDP also increased by a notable 27% in 2018, with the Data Economy – direct and backward indirect impacts - representing 0.22% of Brazil's GDP in 2018).

Table 16: Brazil Indicators - Overview 2016-2018

Brazil – Indicators' Overview						
N.	Name	Metrics	2016	2017	2018	Growth rate 2018/2017
1.1	Number of Data professionals	Total Number of Data professionals (Thousands)	1,160	1,175	1,167	-0.72%
1.2	Data professionals' employment share	% of Data professionals on total employment	1.81%	1.84%	1.82%	-1.42%
2.1	Number of Data Suppliers	Total number of data supplier companies (000s)	35,979	36,906	37,344	1.19%
3.1	Revenues of Data Companies	Total revenues generated by companies specialized in the supply of data-related products and services (Million €)	6,049	6,998	7,592	8.49%
4.1	Value of the Data Market	Estimate of the overall a value of the Data Market (Million €)	6,049	6,998	7,592	8.49%
4.2	Value of the Data Economy (Only Direct and	Direct Impacts (Million €)	6,157	6,996	7,593	8.53%
		Backward Indirect Impacts (Million €)	290	335	352	5.07%

Brazil – Indicators' Overview						
N.	Name	Metrics	2016	2017	2018	Growth rate 2018/2017
	Backward Indirect impacts)					
4.3	Incidence of the Data Economy on GDP (Only direct and backward indirect impacts)	Ratio between value of the Data Economy and GDP (%)	0.16%	0.17%	0.22%	27.45%

Source: European Data Market Monitoring Tool, IDC 2019

Japan

The indicators measuring the state of the Data Market and the Data Economy in Japan have all registered some growth in 2018 over the previous year. The number and employment share of data professionals has risen of 1.6% and 1.3% respectively in 2018, while the number of data suppliers has increased marginally, by 0.5%. This contributed partly to increased growth in revenues and Data Market of 9.7%, but the biggest contributor to this is the increased range of data tools and increased used across most industries of data technologies as opportunities for digital transformation take hold. The Data Economy has also marked a positive development in Japan with an incidence on GDP now at 1.05% of the country's GDP – a growth of 9.8% with respect to the incidence in 2017.

Table 17: Japan Indicators - Overview 2016-2018

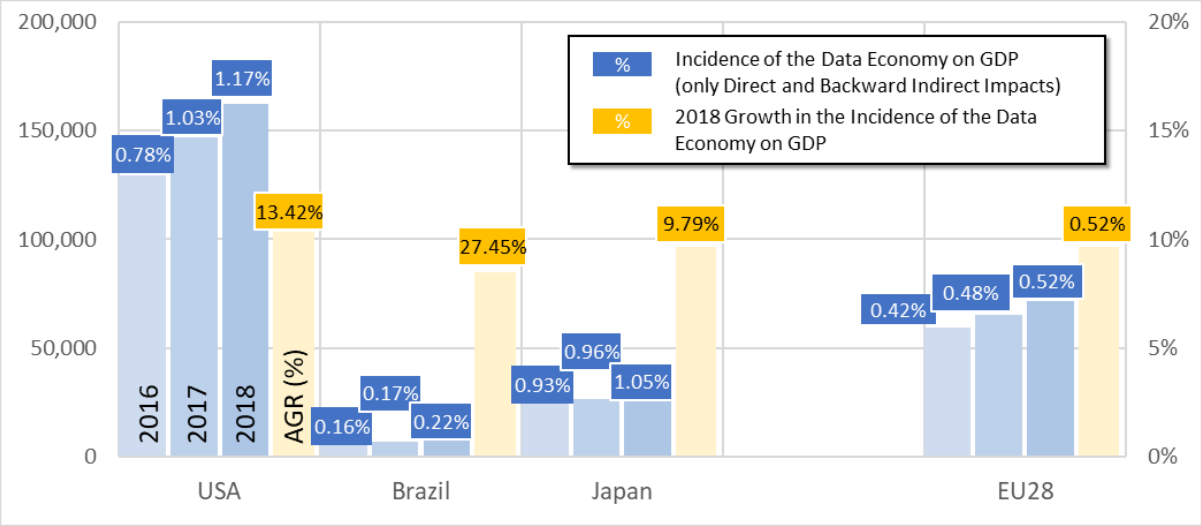
Japan – Indicators' Overview						
N.	Name	Metrics	2016	2017	2018	Growth rate 2018/2017
1.1	Number of Data professionals	Total Number of Data professionals (Thousands)	3,740	4,045	4,111	1.63%
1.2	Data professionals' employment share	% of Data professionals on total employment	5.82%	6.20%	6.28%	1.29%
2.1	Number of Data Suppliers	Total number of data supplier companies (000s)	101,612	104,587	105,103	0.49%
3.1	Revenues of Data Companies	Total revenues generated by companies specialized in the supply of data-related products and services (Million €)	25,513	26,720	29,297	9.65%
4.1	Value of the Data Market	Estimate of the overall a value of the Data Market (Million €)	25,513	26,720	29,297	9.65%
4.2	Value of the Data Economy (Only Direct and Backward Indirect impacts)	Direct Impacts (Million €)	27,394	27,296	29,283	7.28%
		Backward Indirect Impacts (Million €)	1,189	1,230	1,317	7.07%
4.3	Incidence of the Data Economy on GDP (Only direct and backward indirect impacts)	Ratio between value of the Data Economy and GDP (%)	0.93%	0.96%	1.05%	9.79%

Source: European Data Market Monitoring Tool, IDC 2019

International Overview and Comparison with the EU

In line with the results obtained by the previous round of measurement of the international indicators (European Data Market Study Update (SMART 2016/0063), D2.2 First Report on Facts & Figures) the U.S. continue to show the highest impact of the Data Economy on their GDP – 1.17% in 2018, up 13.4% with respect to 2017. While not leading in absolute values, however, Europe emerges as the most dynamic region with a sustained and unsurpassed impacts' growth of more than 9% year-on-year 2017 at the level of the EU27. If the U.K. is added to the equation, the incidence of the Data Economy would be more than 0.5% of the EU28 GDP – a double-digit growth with respect to the previous year. Europe thus presents a growing and dynamic data ecosystem on both fronts – the Data Market and the Data Economy: in terms of size and growth, the value of its Data Market (as defined by the European Data Market Study) is second only to the U.S.; more interestingly, the impact that this market generates on the economy as a whole (the "Data Economy") has become more and more visible over the past few years (2014 through 2017) thus rapidly catching up the gap with the American economy.

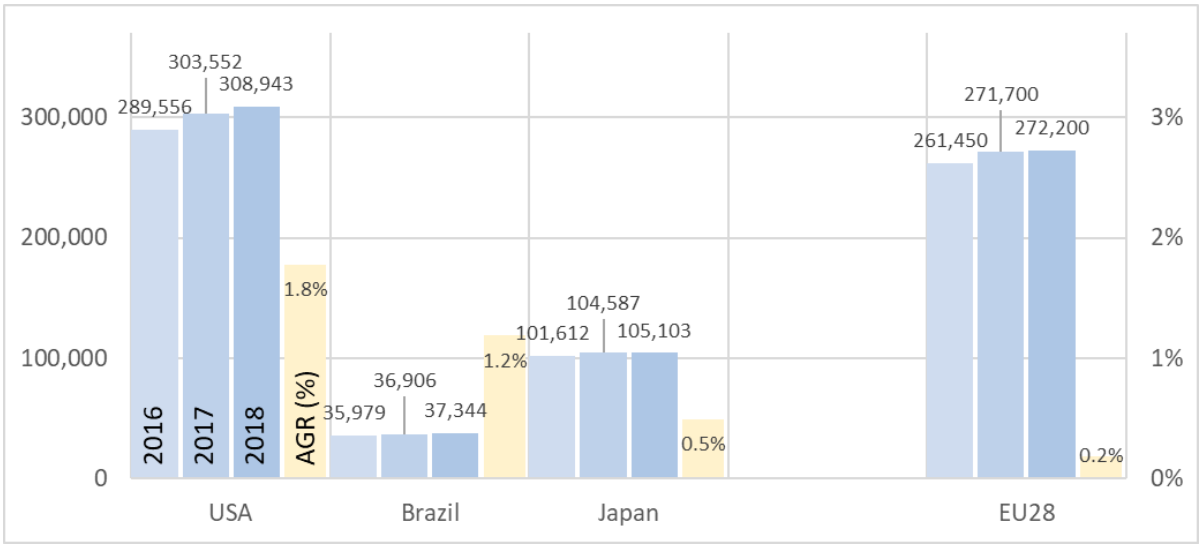
Figure 8: Value, Growth and Incidence of the Data Economy as a Percentage of GDP in the U.S., Brazil, Japan and EU, 2016-2018 (€, Million, %)



Source: European Data Market Monitoring Tool, IDC 2019

Nevertheless, with respect to the number of data suppliers, Europe's growth slows to the lowest of the group, as the U.S. move ahead, in terms of the total number of data suppliers with nearly as many. Europe also easily out-competes Japan though with two companies for every single company in Japan.

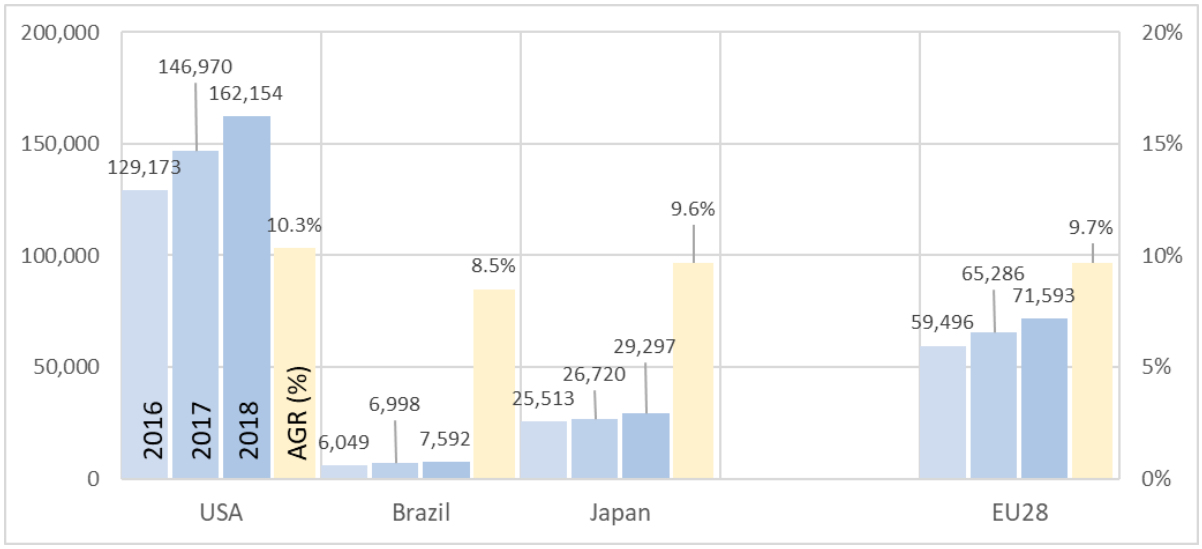
Figure 9: Number of Data Suppliers in the U.S., Brazil, Japan and EU, 2016-2018, Growth 2018 (Units; '000, %)



Source: European Data Market Monitoring Tool, IDC 2019

The Data Market will see the U.S. remaining in the leading position with more than 160 million Euro in size and a buoyant year-on-year growth of 10.3% in 2018 over the previous year. The EU is the only regional market able to challenge the U.S. for dominance in this industry.

Figure 10: Value and Growth of the Data Market in the U.S., Brazil, Japan and EU, 2016-2018, Growth 2018 (Units; '000, %)



Source: European Data Market Monitoring Tool, IDC 2019

3. DESCRIBING THE DATA MARKET – THE QUANTIFIED STORIES

The Second Report on Facts & Figures (D2.4) and the Second Report on Policy Conclusions (D2.5) were accompanied by three quantified stories concentrating on the operational, organisational and/or economic benefits generated by the use of data-driven technologies with a special focus on Big Data and Artificial Intelligence (AI) solutions. While the first story investigated the relationship between the use of data and the means to achieve additional revenues in a business-to-business environment, the second and the third stories focused on the role of data in nurturing and expanding the benefits brought about by AI implementations across a number of different industries and, specifically, on the energy and utilities sector.

3.1 Story 3 - Data Monetization

Constant advances in the process of digital transformation are putting a growing number of companies in the position to use, re-use and exchange data to generate growth. As a result, while still in its infancy, data monetization already constitutes a powerful means to generate additional revenues by using data to add new services to existing offerings, developing new business models, and even directly selling data-based products, services or utilities.

The research identifies three primary paths towards realising that value for data:

1. Direct revenue from data sale/licensing;
2. Additional revenue from bundling data with other services or products;
3. Exchange premiums/trade advantages or discounts.

The first path to data monetization (Type 1) is revenue generated directly from data sales or some form of data licensing. This is notably the case of companies in the information services industry, such as credit bureaus, which collect data about people or organisation and then sell the credit ratings as a service. Type 2 (Additional revenue bundled with other solutions) is obtained when a service provider (e.g. a software provider) combines its "regular" solution with additional value-added services. Exchange premiums or trade discounts (Type 3) do not entail net-new revenue for the data. Instead, something else of value is exchanged.

Through a series of case-studies featuring real-life examples of AI implementations across multiple sectors (from the ICT scene to the Manufacturing industry, from local/regional government to the Telecom sector), the story unveiled a rather complicated situation in Europe in terms of data monetization. As such, the analysis highlighted that:

- **A clear-cut form of data monetization is hard to find**. In fact, a data exchange resulting in the creation of straight revenue streams for data holders as a result of a direct sale (Type 1), or in the form of additional revenue in conjunction with the offering of other services (Type 2), does not appear to be very common in Europe today.
- **Benefits such as operational efficiency, cost optimization, and enhanced quality seem to prevail vis-à-vis data monetization per se**. Indeed, the case studies reveal the prominence of indirect benefits such as operational efficiency, cost optimization, enhanced quality obtained through the data sharing rather than the generation of direct revenue streams.
- **Platforms are the preferred means to perform and enable data monetization**. Buyers and sellers of data rarely connect directly. A technical platform in the form of a data marketplace or a simpler solution such a website is usually preferred to exchange and monetize data.

- **Medium to large companies are at the forefront of data monetization**. While certainly open to all sorts of companies, the services offered by most platforms enabling data sharing, and hence data monetization, appear to be used primarily by medium to large companies.
- **Wide differences in digital maturity, uncertainty surrounding data ownership, and lack of data skills hamper the development of data monetization**. What is more, differences in digital maturity across industries (for example, Financial services and Retail on the one hand and Education and Public sector on the other hand) represent a serious obstacle to the development of data monetization practices.

3.2 Story 4 - How Big Data is driving AI: Selected Examples of AI Applications across European Industries

To better understand how European companies are approaching and implementing AI in a variety of different sectors, the fourth story featured extensive desk research across a multitude of publicly available sources and IDC existing research material. In particular, the story produced five case-studies showcasing how AI solutions currently adopted in Europe to advance an organisation's predictive capabilities allowing for a better interpretation of customer needs and improved forecast accuracy.

The analysis found that AI is embraced to extend the quality of products and services, as well as to obtain overall business benefits in sales & marketing campaigns and customer loyalty programmes. It is fair to say, though, that while AI can be deployed in virtually every sector and for the most disparate reasons, most organisations in Europe are still using AI and BDA to enhance primarily customer experience and satisfaction. A recent IDC's AI survey[3] revealed that 66% of the respondents adopt, or are considering the adoption, of AI systems to increase quality of products and services alongside with creating improvements in customer support. Interestingly, Big Data and Analytics are viewed to serve similar purposes – 44% of the respondents maintain that the main business value driving the adoption of BDA tools and technologies is to improve business process, 43% are using BDA to improve market understanding and 42% indicated product improvement as the third main driver for adoption. Again, a customer-centric approach revolving around better business processes, enhanced products and services, increased market understanding, and improved customer services lies behind the interaction and adoption of AI and BDA.

The case studies presented in this research highlight a number of **immediate benefits** obtained by the organisations adopting AI and Big Data technologies. In essence, they can be summarised as follows:

- **Ease of use and a simplified technology**, freeing data scientists from performing burdensome tasks is a recurring theme in companies such as DAZN (where a very small group of analysts with limited technical background have been able to turn-around the amount of data at their disposal), Coyote (with an increased ease of use of its devices thanks to the adopted technology) and Airbus (with its reduced training time for teams operating the new predictive models).
- **More advanced predictive capabilities,** allowing for a better interpretation of customer needs and improved forecast accuracy is another immediate benefit obtained through the direct application of AI/ML in different case studies, for example Voyage Privé in its ability to better interpret customers' preference signals or AXA and its more advanced predictive model with deeper self-learning capabilities.

[3] AI I Europe: Key Findings of IDC's 2018 AI User Survey; IDC Survey, 2018 (Doc #EMEA44220518, Aug 2018 – IDC Survey) https://www.idc.com/getdoc.jsp?containerId=EMEA44220518

As for the **wider impacts** achieved through the deployment of AI, ML and BDA technologies, they seem to follow a three-way path:

- The first path is the **efficiency obtained by data scientists and the internal users of data** - DAZN for example reports an increase in data scientist productivity of 2.5x with the ability to manage and scale more than 30 data model by a very small and agile team of data scientists. In a similar fashion, Coyote maintains that the application of AI and ML solutions has considerably improved the teamwork and cooperation across its employees.
- The second path is the **impact exerted on the quality of products and services** – AXA claims an increase in forecast reliability of 78%, for example, Airbus a reduction of error rate from 11% to 3% and Coyote an improvement in speed limit detection of 9%.
- The third path revolves around the **overall business benefits** achieved through AI, ML and BDA. Voyage Privé registered a marked improvement in sales & marketing campaigns and a 6% increase in total transaction value per customer; Coyote, again, has been able to strengthen its customer loyalty and marketing efforts claiming an increase of marketing outbound campaign of 11%.

Indeed, AI-Based Automation is shaping across industries and range from task to activity to process level automation balancing human-machine collaboration between producing insights, making decisions and acting on decisions. In another study investigating IT spending forecast in AI and Cognitive Solutions for the period 2018-2022, IDC found that technologies such as mobile solutions, Big Data, cloud, and social media are now 'the new normal' and that the new spotlight is on next-generation technologies. Thanks to their ability to streamline back-office processes and improve customer-facing tasks, artificial intelligence (AI), together with Internet of Things (IoT), blockchain, augmented and virtual reality (ARVR), and robotics is now a key focus for all industries. Some sectors, such as Retail and Financial services, are already more impacted than others. As an example, Sephora – a France-based multinational chain of personal care and beauty stores – implemented AI solutions by operating Facebook messenger chatbots with the aim to increase shop visits and sales. In the U.S. this translated into an increase of 11% in booking rates in 2018 and with an average of US$ 50 (approximately EUR 44) spent by every chatbot driven customer in the same period. In a similar fashion, NatWest - a major retail and commercial bank in the United Kingdom – was able to detect almost than 7 million GBP (more than 8 million EUR) in corporate fraud and avoid substantial related fines thanks to the adoption of AI-powered anti-fraud solutions.

3.3 Story 5 - AI paving the way for the Cognitive Revolution across European Utilities

European utilities are increasingly turning to data-driven technologies to excel in operations, improve customer service satisfaction and create new revenue streams and business models. Indeed, Big Data technologies, smart meter solutions and the increasing deployment of the Internet of Things (IoT), are fuelling an operational and technological revolution in European utilities and leading to the rapid introduction of Artificial Intelligence (AI). In turn, AI is starting to emerge as the technology game changer for the industry, helping utilities make operation more cost effective while ensuring optimal utilization of infrastructure and resources to balance supply and demand safety and reliability.

From fraud detection and claim process automation, to maintenance transformation across solar generation plants and medium voltage power lines, and intelligent handling of customer inquiries, this story investigates a number of real-life AI implementations across a series of European utilities with the aim to unveil some of the most significant quali-quantitative benefits and impacts achieved by the industry in Europe over the past few months.

The desk research and in-depth interviews carried out across the six case studies presented in this research confirm that AI is being successfully employed across a number of European utilities to improve or transform the organisation's operating model, enhance the customer experience, and create new revenue streams and business models. Global services company Atos, for example, is using AI to improve the way utilities fight energy theft (which cost an estimated some €1.3 billion a year to the European gas sector alone). Using AI and automation, Atos customers analysed in this report have managed to increase their ability to identify episodes of energy theft six-fold and reduce the operational cost of their entire revenue protection process by 75%. Enel Green Power, the renewables divisions of Italy-based energy utility Enel, provides another example of the use of AI to transform operations. The company is integrating real-time plant fault recognition technology onboard its maintenance drone fleet in North America, making inspections more efficient and reducing detection-to-repair from days to hour. A case study from E.ON offers further confirmation that asset operations and maintenance are a hotbed for AI. The company has developed analytical models to detect fault predictors for its medium voltage lines and combined these with a self-learning algorithm improving the accuracy of these predictions over time. The result was 100% to 200% better fault prediction after just nine months of use and training of the algorithm. The company estimates that shifting to machine-intelligence based maintenance of network assets provides a 30% improvement, on average, compared to conventional maintenance.

A few take-away points can be highlighted from the analysis of AI implementations in the European utilities sector.

- First of all, the case studies outlined in this research do not represent the mainstream reality for now but their significance – in both "traditional" data-driven use cases (such as those of predictive maintenance, fieldwork optimization and fraud detection) as well as in more innovative realities (such as automated customer service, flexibility, and grid balancing) – is expected to grow remarkedly over the next few years.
- Secondly, the importance of the AI implementation and of the subsequent "cognitive revolution" in the Utilities industry is also destined to increase as a direct consequence of Europe's legacy of early energy market liberalization and leadership in energy system transformation and climate initiatives.
- Thirdly, there is room for policy initiatives at both European and national level as uneven energy market guidelines and regulatory incentives are still at play across the Member States. Some use cases, for example, have been pursued successfully and systematically in some European countries but not in others. To achieve the benefits of a truly European level-playing field and of a truthfully harmonised energy market, more policy guidance would be beneficial.
- Finally, the real-life use cases featured in this research show that the big names in the business and large utilities are among the frontrunners of AI implementations, often co-developing with small companies and startups, and sometimes incubating and providing venture capital at the same time. In addition, the increasingly accurate information produced by AI, and their associated potential on the renewable energy production, are likely to spark higher investments from businesses and augment the willingness of individuals to invest in renewables. For this to concretely accelerate the energy transition towards renewable energy sources and the factual contribution to the EU's long-term strategy for a climate neutral economy, more help (also in the form of funding or financial help otherwise) is likely to be needed, especially for SMEs and individuals.

3.4 What the Quantified Stories tell us so far

An overarching common theme seems to underlie the three quantified stories that accompanied the Second Report on Facts and Figures and the Second Policy Report – the need to obtain accurate, reliable and vast amount of data so that:

a) data can be safely shared and increasingly monetised; and
b) data can effectively sustain and nurture the expansion of AI technology implementations.

The in-depth interviews that were conducted for the realisation of the research on data monetization in Europe revealed that data monetization appears to be a still unacknowledged and somewhat underestimated phenomenon. Data holder and supplier companies need to become more aware of the opportunities offered by their data in terms of new revenue generation, additional income obtained in conjunction with the offering of other solutions and possible premiums or other forms of business advantages that can be achieved by sharing their data with other organisations. Data re-users, on the other hand, need to better understand the advantages of acquiring data from other companies for their internal usage (to improve productivity, better manage costs, and enhance customer relationships) or with the aim of creating brand new business opportunities. Also, the need to create a trusted environment around data monetization emerged as an important theme for data suppliers and data users so to enable data exchanges and allow for subsequent opportunities of data monetization. Uncertainties about the regulatory environment around data sharing and data monetization were also indicated as a possible barrier to further development of data monetization.

To counteract these barriers, an increase in funding for SMEs allowing them to engage in data monetization emerged as an important factor to sustain data monetization exchanges, since effective data monetization requires secure and first-rate technical solutions, advanced technical skills, as well as business acumen and the ability to understand how to use or re-use data in an increasing digitised economy. Additional research into the theme of data monetization could also be beneficial as data monetisation may take several forms, evolve towards additional practices not yet known, and provide extra benefits not yet considered. More research on data monetization mechanisms at play in other parts of the world, especially in the United States, Japan but also in China and India, could therefore help the European Data Economy grasp new opportunities related to data sharing and monetization.

Reliable, trustable, secure and accurate data can not only be safely shared and increasingly monetized. They can also effectively sustain the successful deployment of AI technologies. Whether in the form of more internal efficiency for analysts and data scientist teams, improved quality of products and services, or wider business benefits in terms of sales and marketing, Big Data technologies appear to be instrumental in providing seamlessly large and accurate amount of data to AI technologies, which, in turn, are of paramount importance in bringing about concrete and quantifiable benefits in different industry sectors – from Manufacturing to Telecommunications, from local and regional government to the Energy sector and the Utilities. The present AI expansion, as well as its significant growth potential expected over the next few years, is directly dependent on the availability of huge data sets combined with increases in computing power, connectivity, and data-related technologies. Further action to facilitate the sharing of data held by public and private sectors in Europe, as well as a coordinated strategy in Research & Innovation investments to ease access to connectivity, interoperability, aggregation of public data, or to support the development and implementation of a data infrastructure to enable the management and sharing of data in real-

time should be fostered at EU level, as well as at national level in coordination with the European Commission.

4. MAPPING THE DATA MARKET – DATA LANDSCAPE AND DATA MARKET MONITORING TOOL

4.1 The EU Data Landscape

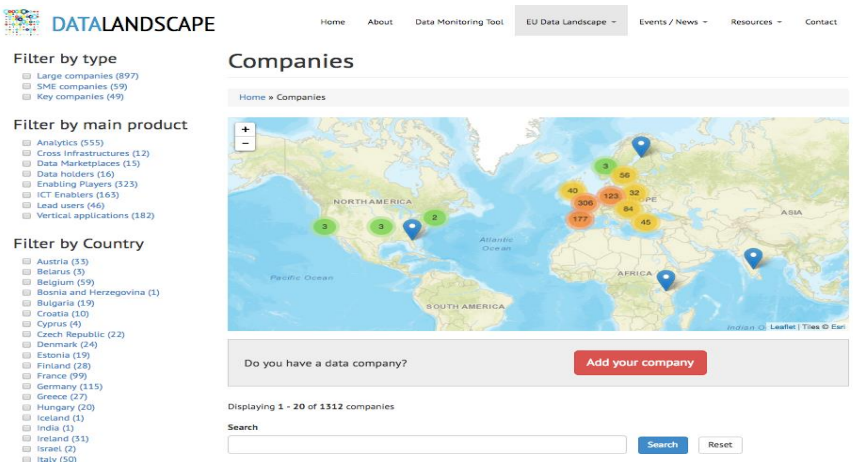
The Second EU Data Landscape Report (D4.2) provides an overview of the EU Data Landscape database revision performed by the study team between January 2018 and January 2019. With a total of 1459 companies and coverage of 41 countries (European Union-28, Belarus, Bosnia and Herzegovina, Georgia, Iceland, Israel, Kenya, Moldova, Norway, Serbia, Switzerland, Turkey, Ukraine and the United States), the database offers a comprehensive overview of the most important data companies in Europe. More specifically, among the 1459 companies, 259 have been identified as *Key Data Landscape companies* in line with a set of criteria adopted and further specified in the following paragraphs.

Main changes to the EU Data Landscape

The 2019 EU Data Landscape review introduced some changes to the approach:

- This year's research was widened by the use of a revised methodology and more efficient sources, with a specific focus on Key Data Landscape Companies.
- The existing EU Data Landscape database was validated and further extended in geographical and coverage scope. In particular, as regards geographical coverage, the database currently includes companies from 41 countries, against the 36 of the January 2018 update.
- The list of Key Data Landscape companies was reviewed according to updated or new criteria identified by the study team, leading to 150 new entries in this category.

Figure 11: Database of Data Landscape companies (www.datalandscape.eu)



Source: http://datalandscape.eu/companies

Overview of the EU Data Landscape Database (status in January 2019):

- Overall, the database has grown by 16% from 2018 to 2019 with the addition of 203 new companies. Out of the 203 new companies, 165 are Key Data Landscape companies.
- UK companies account for 24.4% of the total database, followed by Spain (13,2%), Germany (8,8%), France (8,6%), Netherlands (4,6 %) and Belgium (4,1%).

- Analytics continues to be the most populated category (accounting for 44% of the database).
- The share of companies categorised as Vertical Applications grew by three percentage points reaching 17% of the database (from 169 companies in 2018 up to 246 in 2019).
- In the AI and Machine Learning category, the UK accounts for 25% of the companies, followed by Spain (12%), Germany (11%) and France (11%).

Key Data Landscape companies (status in January 2019):

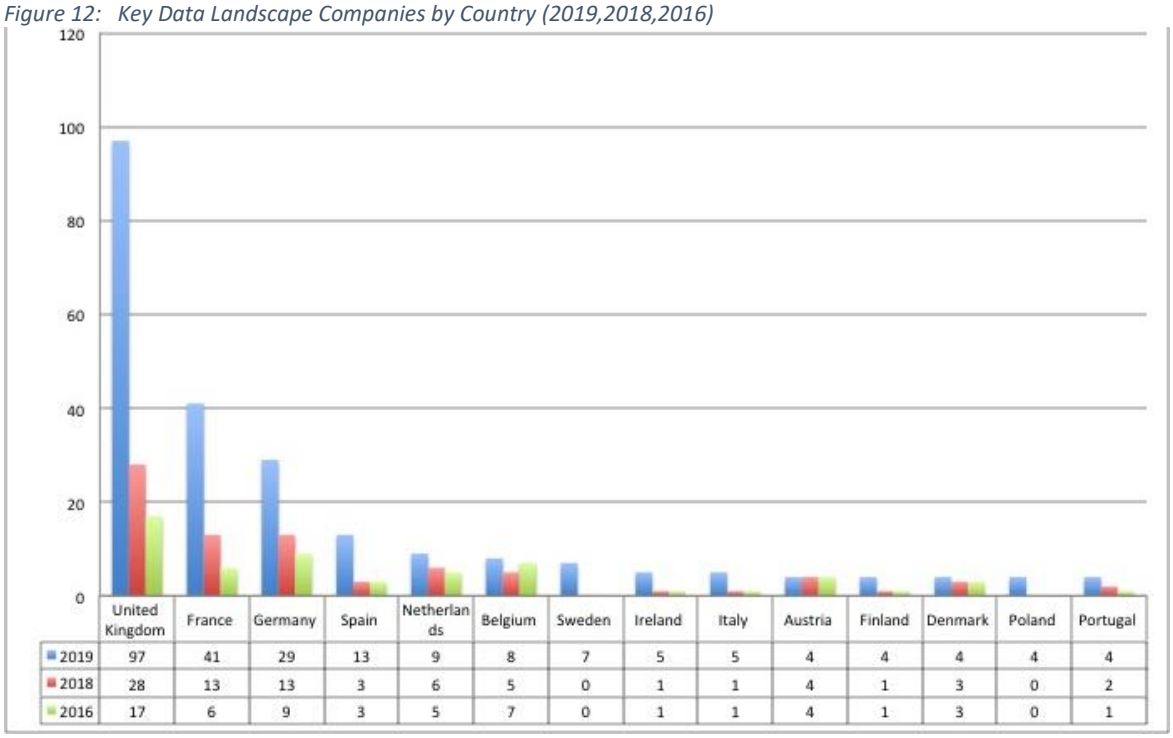
- Focusing on the methodological approach, Key Data Landscape companies were selected from the main database according to the following pre-set criteria:

    - The company is listed in the Global Big Data Landscape map,[4] or
    - The company and its proof of concept is already established enough – as a proxy the study team took into account companies receiving over EUR 1m in funding according to Crunchbase database, which provides data on the world's most innovative companies, including data on the amount of capital obtained, and
    - The company has its main headquarter or R&D department in Europe.

- The list of Key Data Landscape companies grew from 94 in 2018 to 259 in 2019. The 2019 update of EU Data Landscape put special focus on the Key Data Landscape Companies category, adding 150 new entries. For this purpose, the main source used was Crunchbase's list of EU Big Data Companies.
- Most Key Data Landscape companies are headquartered in the UK (97, 37.5%), followed by France (41, 16%) and Germany (29, 11%).
- In 2019 *Analytics* has become the predominant category (49%), overcoming *Vertical applications* (following with 43%).

[4] Matt Turck (2017). Big Data Landscape 2017, Firstmark. Available at: http://mattturck.com/wp-content/uploads/2017/04/Big-Data-Landscape-2017-Matt-Turck-FirstMark.png

Figure 12: Key Data Landscape Companies by Country (2019,2018,2016)



	United Kingdom	France	Germany	Spain	Netherlands	Belgium	Sweden	Ireland	Italy	Austria	Finland	Denmark	Poland	Portugal
2019	97	41	29	13	9	8	7	5	5	4	4	4	4	4
2018	28	13	13	3	6	5	0	1	1	4	1	3	0	2
2016	17	6	9	3	5	7	0	1	1	4	1	3	0	1

Source: European Data Market Study, D4.2 EU Data Landscape, Review at January 2019

4.2 The European Data Market Monitoring Tool

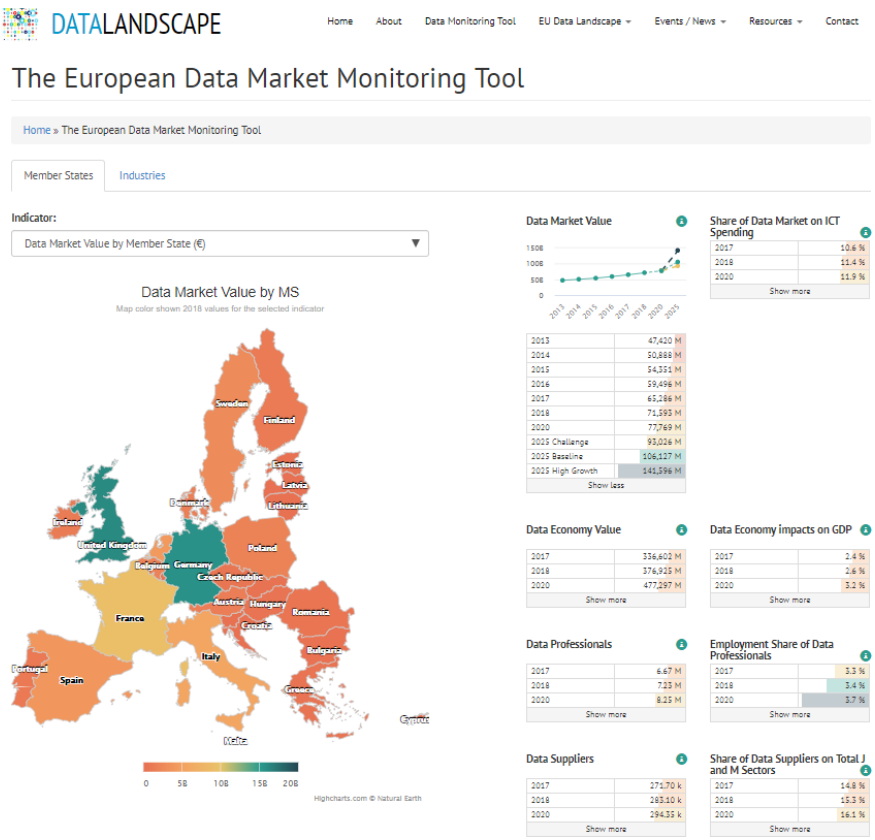
Along with the validation and extension of the EU Data Landscape database presented in the previous paragraphs, the study team performed an update of the European Data Market Monitoring Tool launched on www.datalandscape.eu in April 2018.

Based on the Dataset accompanying the Second Report on Facts and Figures (D2.4), offering an overview of the data based on which the report was drafted, the scope of the data showcased by the EDM Monitoring Tool was extended to include the key facts and figures measured for the year 2018 at the level of the total EU28 and for all 28 EU Member States, when available and applicable. Industry-specific views were also offered with indicators provided by industry sector when possible.

The figures presented for the years 2020 and 2025 (three scenarios) were also updated to reflect the results obtained from the second round of measurement of the EDM indicators.

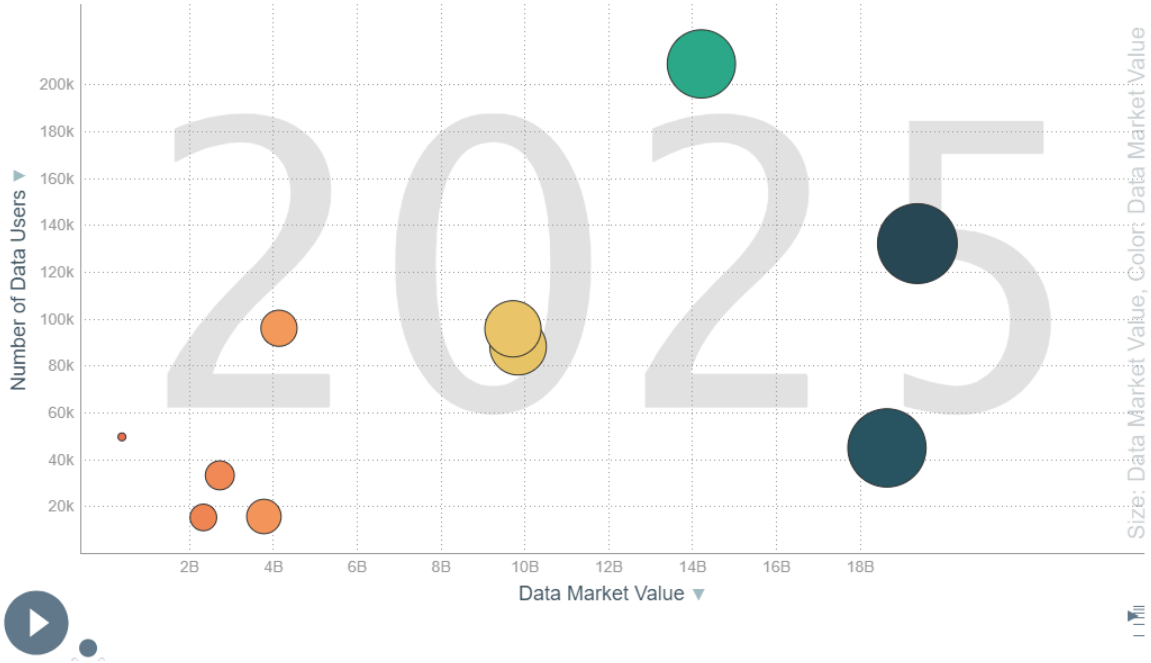
An overview of the different sections of the European Data Market Monitoring Tool is provided in the figures below.

Figure 13: The European Data Market Monitoring Tool – Member States



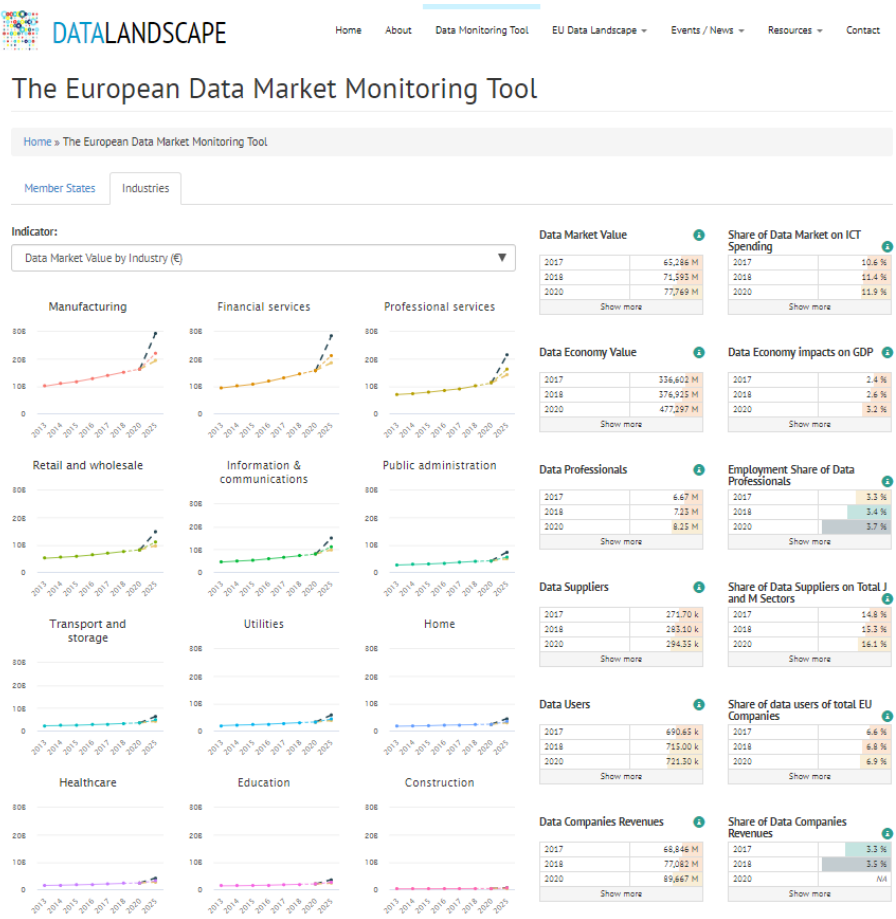
Source: http://datalandscape.eu/european-data-market-monitoring-tool-2018

Figure 14: The European Data Market Monitoring Tool – Bubble chart



Source: http://datalandscape.eu/european-data-market-monitoring-tool-2018

Figure 15: The European Data Market Monitoring Tool – Industries



Source: http://datalandscape.eu/european-data-market-monitoring-tool-2018

5. ACTING UPON THE DATA MARKET – THE ROLE OF POLICY

By investigating the role of policies in shaping the present and future trends of the European Data Market and Data Economy, the Second Report on Policy Conclusions (D2.5) complemented the sizing and forecasting exercise carried out for the Second Report on Facts and Figures (D2.4), as well as the additional analysis obtained through the three quantified stories.

5.1 The Role of Policy and the Future of Europe's Data Economy: The Three Scenarios

The Second Report on Policy Conclusions (D2.5) presents the alternative evolution paths of the European Data Market and Data Economy at 2025, described as three potential scenarios driven by different macroeconomic and framework conditions, shaped by critical turning points to be faced in the next years by governments, businesses and social actors. The scenarios are an update of those presented in March 2018, building on the updated EDM dataset and forecasts and insights from last year's events. The 2025 scenarios main focal issues are confirmed as following:

- **The Data Market's pace of growth:** how fast will data-driven innovation grow in Europe? The scenarios outline the three alternative possibilities of slow, medium or fast pace of innovation, **including the speed of AI solutions take-up**, which will clearly be a differentiating factor, given the interdependence between the availability of high-quality data and AI services and applications.
- **The potential evolution of the model of data governance**, in terms of how the ownership, access, control and exploitation of data assets will be managed. To put it more bluntly: who will have power on data and what will governments do about it? The scenarios outline future models ranging between two potential extremes: on the one hand, a data governance model where a few data holders (private or public) control most of data assets; on the other hand, an open and participatory data governance model, based on sharing and transparency. **Using the terminology developed in the last year for AI, this should be the trustworthy and human-centric Data Economy Europe aspires to.**

The Data Economy scenarios therefore are positioned at the intersection of these two main focal issues as follows:

- The **Baseline scenario** is characterised by a healthy growth of data innovation, a moderate concentration of power by dominant data owners with a data governance model protecting personal data rights, and an uneven but rather wide distribution of data innovation benefits in the society. This is considered the most likely scenario.
- The **High Growth scenario** (Data-driven reality) is characterised by a high level of data innovation, low data power concentration, an open and transparent data governance model with high data sharing, and a wide distribution of the benefits of data innovation in the society;
- The **Challenge scenario** (Digital Maze) is characterised by a low level of data innovation, a moderate level of data power concentration due to digital markets fragmentation, and an uneven distribution of data innovation benefits in the society.

The scenarios explore the drivers and framework conditions which may lead to maximise the benefits of a balanced Data Economy and to avoid the risks of an unbalanced one, highlighting the consequences of policy actions.

Policy and the Baseline Scenario

The Baseline scenario at 2025 predicts a healthy growth of data-driven innovation, supported by increased investments coming from large advanced enterprises and a minority of competitive SMEs savvy in the use of ICTs. In this framework, leading companies and regions are expected to increase their advantage, while more traditional companies and sectors struggle to move at the same speed, especially where 5G infrastructures needed to support next generation online services, especially IoT, are not widely diffused.

The EU GDP cumulative growth average in the period 2018-2025 (+1.7%) is expected to sustain the investments in the digital economy and consumer willingness to spend. Enterprises will accelerate their digital transformation process, enabled by data-centric processes and new digital business models. This acceleration will lead companies to add more than 3 million data professionals' positions between 2020 and 2025, which will not only increase the data professionals skills gap but also trigger the development of potential data talent "wars" for the most skilled professionals.

In this Baseline scenario, policy is expected to reach a partial success. In 2019 the effects of some policy measures aiming at building the Digital Single Market (DSM), first of all the General Data Protection Regulation (GDPR) and the Free Flow of non-personal Data Regulation (FFoD), are starting to be seen. While GDPR has set itself as a worldwide standard for privacy protection, gaining attention and being emulated also in the U.S., adjustments and revisions linked to implementations hurdles are also expected to come. The so called GAFA (Google, Facebook, Amazon, Apple) will maintain their dominant position a regards personal data flows but will need to become more transparent and guarantee users more control on how their personal data is used and shared.

At the same time, the entry into force of the FFoD Regulation will be essential in order to unlock the exploitation of European datasets and enable new data-driven processes such as Machine Learning. The completion of the DSM is expected to progress and possibly be completed only at the end of the forecast period. In this scenario, European investments in digital innovation will increase strongly through the Horizon Europe and the new Digital Europe Programme, with support by industry through the Public-Private Partnerships (PPP) and their collaborations (currently there are agreements between BDVA (Big Data Value Association) – ETP4HPC (European Technology Platform for High Performance Computing), AIOTI (The Alliance for Internet of Things Innovation), and expectations for a new AI PPP).

Policy and the High Growth Scenario

According to our latest estimates related to the High Growth scenario, the growth of the European Data Market will enter a faster growth trajectory than in the Baseline scenario and the adoption of data technologies will spread beyond the minority of pioneers to a wider population of mainstream users. Digital transformation, data monetization, B2B data sharing on multi-user cloud platforms will spread faster than in the Baseline scenario. All industries will keep pace, also the Public sector, even though the intensity of data innovation will grow faster in Finance, Manufacturing, Professional services, ICT and Media. The supply-demand dynamics will change from technology-push to demand pull, with a fully developed ecosystem generating positive feed-back loops between data companies and users.

In this scenario, the EU GDP CAGR in the period 2020-2025 (+2.2%) will be 2.5 times higher than in the Challenge scenario and 1.5 times higher than in the Baseline scenario. This will accelerate the investments in the digital economy and consumer willingness to spend. In the European Union public and private investments will accelerate in Artificial Intelligence, advanced robotics,

automation as well as new skills. In fact, the potential data professionals skills gap will grow exponentially to 1.5 million unfilled positions in the EU27 and 1.7 million unfilled positions in the EU28: this risk will need to be managed in advance or the lack of skills will become a serious constraint for data-driven companies and data suppliers.

Policy measures at European level have a relevant role to play in this scenario. The completion of the Digital Single Market within the forecast period is a key success factor, requiring rapid and successful implementation of the GDPR and the Free-Flow of non-personal data initiative, with true liberalisation of data flows across Europe. At the same time, R&D investment at EU and national level must be particularly effective and the Network of Excellence Centres must succeed in becoming innovation beacons in every region as well as increasing the supply of scarce data skills. In short, in this scenario Europe captures in full the digital opportunity.

Policy and the Challenge Scenario

The Challenge scenario at 2025 represents a more extreme version of the Baseline scenario, where markets fragmentation and the failure to complete the Digital Single Market exacerbate regional and industry differences, so that only the best enterprises and the richest regions keep pace with the technology race. European innovation forces become lost in a maze of barriers and are unable to overcome them, resulting in a much slower Data Market and Data Economy growth. This scenario foresees a negative self-reinforcing circle, where less positive global economic conditions discourage investments and weaken global demand with a negative impact on European growth. A slower pace of digital innovation deprives the economy of the boost to growth potentially given by data-driven services and products, while enterprises find competing in international markets more difficult.

In this Challenge scenario, the EU GDP CAGR in the period 2020-2025 will be only 0.9%, substantially lower than in the other scenarios. This could be potentially caused by relevant geo-political shocks in critical regions, trade wars fostered by new protectionist policies; or else a new financial crisis. In this scenario the economy would also suffer from an unexpected resistance to digital transformation by enterprises and the population due to reluctance to change, and difficulty in adopting new processes and ways of working, compounded by a relatively old working population and lack of specialist skills.

Concerning the role of policies, this scenario is driven as much by the failure of the Digital Single Market and of innovation investments than by global economic trends. The risk of insufficient digital capacities and networking powering data innovation across Europe would create a disadvantage for the European industry and a growth gap with other world regions, particularly Asia-pacific.

5.2 Value from Data, the new Priority of Data Users

According to our latest estimates, the European Data Market has increased by over 50% from 2014 to 2018 (from €47B to €72B in the EU28), which represents a remarkable increase. But even more remarkable has been the change in awareness about the value of data and the potential of data-driven innovation to increase growth and welfare. In fact, only in the last couple of years the role of data as the key ingredient to capture value from digital innovation has become clearly recognised. Collecting, analysing and exploiting data is understood as the enabling condition of digital transformation, as technologies such as the Internet of Things, Virtual/Augmented Reality, Cybersecurity, Robotics, Additive manufacturing (3D printing) rely on the collection and manipulation of data to be implemented. While in the early phase of development of the data market enterprises focused on the collection and management of data, now business users prioritize strongly the extraction of value from data, which includes also data monetization.

In this framework, the connection between Big Data and Artificial Intelligence is establishing itself as the real game changer. Big Data is in fact the fuel powering the emerging AI innovation wave. The European Union has recognised the relevance of the forthcoming AI revolution by launching an AI Communication and Action Plan, coordinating efforts and increasing investments to make sure that Europe keeps pace with international competition. At the same time, Europe has chosen to deal also with the potential risks of Artificial Intelligence, aiming for a human-centered AI.

This year 2019 will also see the completion of the mandate of the current European Commission and Parliament, with a new Commission to be appointed by the end of the year. Nevertheless, the need for Europe to capture the digital opportunity will remain at the forefront of European strategies. Next year, European Parliament and the Council will finalise the new Multi-Annual Financial Framework 2021-2027 proposed by the Commission making decisions which will influence the size and directions of the investments in digital technologies and policies. In this scenario, market intelligence is needed more than ever to inform and support strategic decisions.

5.3 The EU Data Policy and the International Dimension

The international indicators emerging from the second round of measurements of the European Data Market Monitoring Tool continue to highlight the prominence of the U.S. in terms of size of the Data Market and the Data Economy, even if Europe has undertaken significant steps in 2018 towards a more coordinated approach and a progressive removal of the barriers identified by the Digital Single Market Strategy. The EU28 continues to represent the second-largest Data Economy worldwide after the U.S., followed by Japan and Brazil, even if it lags behind both the U.S. and Japan with regards to the incidence of the Data Economy in terms of direct and backward indirect impacts on GDP. Furthermore, Europe lags behind in terms of digital infrastructure and number of data professionals. Not only does it suffer from high levels of fragmentation when it comes to the use of digital technologies across companies, but it is also affected by the lack of the big tech corporations, which represent two of the main reasons linked to Europe's relative delay.

Filling this gap would be essential to increase Europe's competitiveness and for the future of work in the EU. More specifically, upskilling will be crucial as technological advances in AI, robotics, IoT, 3D printing and quantum computing, just to name a few, are having a very significant transformative effect on work, triggering improved quality and job satisfaction but also implying potentially disruptive social challenges, which need to be tackled at EU level.

As regards the progressive removal of the challenges identified by the DSM Strategy, noteworthy policy measures in 2018 in this respect have been the entry into force of the General Data Protection Regulation, the Regulation on the free flow of non-personal data, as well as the Regulation addressing unjustified geo-blocking. In 2019 other important steps have been made through the European Parliament's and Council of the European Union's approval of the new Copyright Directive during last March and April, while intensified efforts towards the adoption of an ePrivacy Regulation, necessary to complete the EU's framework for data protection and confidentiality of communications, will probably foster discussion in the next period.

Europe's competitiveness at international level will benefit from the development of a common data space where European companies would be supported by increased data sharing in the scientific and private sector, but also through accessibility and re-use of public and publicly funded data. The re-use of these data can in fact not only contribute to the growth of the European economy, but also to the development of AI and the fight against societal challenges. It is exactly in the field of AI that

Europe has recently launched new policy initiatives[5] aimed at filling the gap with other world regions, and in particular the U.S., as far as innovation and R&D are concerned.

Within this framework, Europe's efforts in setting standards not only as far as ethics principles are concerned, but also as regards data protection, privacy and fair competition, have led to greater attention towards these topics on an international level and can put the grounds for Europe to be a role model for other countries.

[5] Including the European Strategy on AI from April 2018, the Coordinated Plan on Artificial Intelligence released in December 2018 and the EC appointment of the High-Level Expert Group on AI working on the development ethics guidelines for trustworthy Artificial Intelligence.

6. CONCLUSIONS

6.1 Quantifying the European Data Market – Key Facts & Figures

The year 2018 sees all the indicators measured by the EDM Monitoring Tool in a positive dynamic from 2017, as the European economy continues its development cycle. The value of the Data Economy, which measures the overall impacts of the Data Market on the economy as a whole, is to exceed the threshold of 300 Billion Euro in 2018 for EU28, with a growth of nearly 12% over the previous year. The positive trend in the growth of the Data Economy is confirmed by the EU28 Data Market value in 2018, which is displaying a growth rate well above the one exhibited by the total IT spending, at 9.7% year-on-year, reaching 71.5 Billion Euro.

The EDM Monitoring Tool has been analysed along four main dimensions:

- The Workforce and Skills dimension - including the measurement of data professionals and their potential skill gap.
- The Supply and Demand dimension - incorporating the measurement of data supplier and data user companies and the revenues generated by these companies.
- The Business and Economy dimension - comprehending the size of the Data Market and the value of the Data Economy.
- The International context dimension - including a select number of indicators for Brazil, Japan and the US.

Figure 16: The four Dimensions of the Data Market's Key Facts & Figures

Source: The European Data Market Monitoring Tool, IDC, 2019

The Workforce Dimension: Data Professionals and Data Professionals Skills Gap

Data professionals are estimated at 5.7 million in the EU27 and 7.2 million in the EU28 in 2018, thus marking a significant increase in 2018 over the previous year (8.6% and 8.4% year-on-year respectively). By 2020 data professionals are estimated at 6.6 million in EU27, and 8.3 million in EU28 (registering a Compound Annual Growth Rate with respect to 2018 of 7.2% and 6.9% at the level of EU27 and EU28 respectively). This increase confirms the positive evolution of the workforce involved in data-related professions over the period under consideration. Furthermore, data professionals will continue to rise in the long period, accounting for more than 13 million units in EU28 and almost 11 million data professionals in EU27, according to our estimates. The number is

expected to rise at a CAGR of 6.6% in EU28. This growth could be as high as 9.8% per year if the High Growth scenario is achieved.

As regards data professionals' skills, the Data Professionals Skills Gap indicator has highlighted an imbalance between demand and supply of data skills in Europe since the first measurement for the year 2014. In the year 2018 the strong increase of demand of data professionals continued (+7%), the estimated gap grew by 10% reaching approximately 571,000 unfilled positions in the EU28 (496,000 without the U.K.), corresponding to 7.2% of total demand (7.9% without the U.K.) By 2020 the gap is expected to expand to 641,000 unfilled positions in EU28, corresponding to 7.8% of total demand (9.1% without the UK, where slower growth is expected due to the impacts of Brexit).

The three forecast scenarios at 2025 all predict an increasing data skills gap: the forecast gap ranges from 9% of demand in the Challenge one, to 10% in the Baseline, to 14% in the High Growth (EU27). The slightly lower dynamic of data skills demand in the U.K. (due to slightly slower forecast of the Data Market growth) means that the gap share of demand is smaller in EU28 than in EU27.
The absolute size of the data skills gap is relevant, potentially reaching 925,000 unfilled positions in 2025 in the EU27 Baseline scenario, but up to over 1.5 million in the EU27 High Growth scenario. In the Challenge scenario the data skills gap is forecast at 775,000 unfilled positions in 2025. This underlines the need for policy action to prevent and minimize the unbalance between data skills demand and supply in the next years.

The Supply - Demand Dimension: The Data Companies

According to the second round of measurement of the EDM indicators, the number of data suppliers continue to grow at a faster pace than the numbers of data users in the longer term – out to 2025. Data suppliers are estimated at almost 145,100 in the EU27 and at more than 283,000 units in the EU28 for 2018, thus exhibiting a year-on-year growth of 4.1% and 4.2% respectively. Data users, instead, are projected to grow at 3.4% in 2018, amounting to more than 534,000 in the EU27 and to 715,000 units in the EU28. If compared to the measurements carried out by the European Data Market Monitoring Tool over the period 2013-2015, these latest estimates show a picture of consolidation of data companies in the EU, following increasing growth rates increasing over the prior four years.

Revenues generated by data suppliers have registered a constant increase over the past three years to reach nearly 59 Billion Euro in EU27 and 77 Billion Euro in EU28. Data companies' revenues account for 3.5% of total company revenues in 2018. Data companies' revenues are expected to follow the Data Market, as imports and exports of data tools and services tend to follow each other. Forecasting data companies' revenues shows an expected annual growth rate out to 2025 of 7.1% - easily outpacing the growth of the total ICT market over the same period (expected to be 1.5% from 2020 to 2025 Baseline). The smaller Member States show the highest long-term growth as they have a smaller base from which to grow, but the larger Member States will make the biggest overall contribution to the Data Economy out to 2025.

The Business and Economic Dimension: The Data Market and the Data Economy

The value of the European Data Market is expected to reach 77.8 Billion Euro, with a growth rate of 97% in 2018, and at an average rate of 4.2% out to 2020. Each of the Member States shows strong growth, well ahead of the expected growth for the ICT market as a whole, which is expected to grow only by 1.9% in 2018, and at an annual rate of 2.0% to 2020. The Data Market share of total ICT is 11.4% for 2018 and will grow out to 2025.

The larger industries, accounting for the greatest number of companies, represent the largest share of the Data Market. In terms of adoption by industry, the highest rates of Data Technology tend to be in Finance, Professional services, and in Retail. Thanks to their size, these industries are the biggest consumer of data technologies. Manufacturing's sheer size in the EU economy makes it the largest industry in the Data Market. However, there is significant scope for increased adoption of data technology in the manufacturing industry, so its leading position is unlikely to change.

The Data Market will continue to out-grow the total ICT market, with its share of this market rising from close to 11% in 2016, to more than 15% by 2025. The forecast for the Data Market shows which industries make the biggest contribution to the overall market growth, and the key industries of Manufacturing, Finance, Professional services, Information Technology, and Retail account for more than 75% of the total market growth from 2018 to 2025, with nearly 60% coming from the three main industries of Manufacturing, Finance, and Professional services.

The value of the Data Economy for EU28 has been estimated to exceed the threshold of 300 Billion Euro in 2018, overall confirming the estimated growth of nearly 12% of the previous round of measurement in February 2018.

The overall impacts for EU27 and EU28 will be 4.0% and 4.2% of the GDP respectively in the Baseline scenario by 2025. Indirect and induced impacts will continue to be evenly distributed in 2018. As in the previous round of measurement of the European Data Market Monitoring Tool in February 2018, data user companies will continue to consolidate the quantitative benefits stemming from the use of data thus contributing to the importance of indirect impacts. Not surprisingly, these benefits will go beyond the users and will translate in higher induced effects, generating jobs and revenues beyond the data companies itself. The positive conditions under the High Growth scenario will lead the overall impacts exceed 1000 Billion Euro in 2025. In the High Growth scenario, the penetration rate of data products and services into the user industry will be lower with respect to the other scenarios and replaced by higher induced effects.

A screenshot of the Data Economy by industry shows that the Financial sector, the Manufacturing industry and the realm of Professional services continue to represent the vertical markets in which the impacts of the Data Economy are most strongly felt. Thanks to the significant diffusion of data-related technologies, these industries exhibit high levels of forward and backward impacts and can convey effects at an induced level more quickly and more effectively than other industries. Their IoT diffusion and the usage of Cloud Computing, as well as the usage of mobile and social technologies, coupled with the ongoing process of digital transformation, make these industries particularly reactive to induced effects. Emerging technologies such as Artificial Intelligence and blockchain applications, are also gaining momentum in these industries, thus reinforcing the impact of indirect and induced impacts in these sectors.

The International Dimension - The Data Economy Beyond the EU – US, Brazil and Japan

The latest data stemming from our international measurements shows that the European Data Market and economy in the period 2017-2018 continues to consistently hold second place after the U.S. in value but slips to third in growth. Europe still presents a growing and dynamic data ecosystem on both fronts – the Data Market and the Data Economy: however, it lags both the U.S. and Japan in terms of the incidence of the Data Economy on GDP and has some catching up to do.

The positive development of the U.S.' Data Economy is confirmed by a solid year-on-year growth of the main indicators monitored, including the number of data professionals, companies, and the overall Data Market.

Brazil confirms a lower growth rate following another difficult year for the country's general economy. However, it shows signs of recovery on some key areas of the Data Economy and still has the potential for significant growth over the next few years.

Japan's Data Market is the similar to the European one in terms of growth and investment, but still only half the size. It competes with the EU across data professionals and data suppliers, but its growth is significantly lower than for Europe. The economy continues to suffer from weakening internal demand and lack of consumption, even though it showed a small lift in Q4 2018. This, in turn, slows IT spending and, as a result, limits the Data Economy and the Data Market potential.

Looking at the estimates of the data suppliers, the EU exhibits a year-on-year growth 2017-2018 of 0.2% - notably lower than the U.S., which showed 1.8% growth over the same period for the number of data suppliers.

6.2 Describing the Data Market – The Quantified Stories

The three quantified stories produced by the study team between August 2018 and May 2019 focused on the operational, organizational and/or economic benefits generated by the use of data-driven technologies with a special focus on Big Data and Artificial Intelligence (AI) solutions.

Constant advances in the process of digital transformation are putting a growing number of companies in the position to use, re-use and exchange data to generate growth. As a result, data monetization already constitutes a powerful means to generate additional revenues by using data to add new services to existing offerings, developing new business models, and even directly selling data-based products, services or utilities. Reliable, trustable, secure and accurate data can not only be safely shared and increasingly monetised. They can also effectively sustain the successful deployment of AI technologies. Whether in the form of more internal efficiency for analysts and data scientist teams, improved quality of products and services, or wider business benefits in terms of sales and marketing, Big Data technologies appear to be instrumental in providing seamlessly large and accurate amount of data to AI technologies, which, in turn, are of paramount importance in bringing about concrete and quantifiable benefits in different industry sectors – from manufacturing to telecommunications, from local and regional government to the energy sector and the utilities. The present AI expansion, as well as its significant growth potential, which is expected over the next few years, is directly dependent on the availability of huge data sets combined with increases in computing power, connectivity, and data-related technologies. Further action to facilitate the sharing of data held by public and private sectors in Europe, as well as a coordinated strategy in Research & Innovation investments to ease access to connectivity, interoperability, aggregation of public data, or to support the development and implementation of a data infrastructure to enable the management and sharing of data in real-time should be fostered at EU level, as well as at national level in coordination with the European Commission.

6.3 Mapping the Data Market – Data Landscape and Data Market Monitoring Tool

The Second EU Data Landscape Report (D4.2) offered an overview of the EU Data Landscape database revision performed by the study team between January 2018 and January 2019. With a total of 1459 companies and coverage of 41 countries, the database offers a comprehensive overview of the most important data companies in Europe. Overall, the database has grown by 16%

from 2018 to 2019 with the addition of 203 new companies, among which 165 are Key Data Landscape ones.

Moreover, the study team performed an update of the European Data Market Monitoring Tool launched on www.datalandscape.eu in April 2018.The scope of the data showcased by the EDM Monitoring Tool was extended to include the key facts and figures measured for the year 2018 at the level of the total EU28 and for all 28 EU Member States, when available and applicable.

6.4 Acting Upon the Data Market – The Role of Policy

The second round of measurement of the European Data Market Monitoring Tool reveals a positive picture of Europe's data market and data economy with most of the indicators measuring the development of the data ecosystem in Europe steadily on the increase. In 2018, AI conquered a stable position at the top of the European policy agenda with the launch of an ambitious strategy (Communication on AI, April 2018) and a comprehensive Action Plan (December 2018). This strategy has been endowed with additional instruments at the EU level, such as an AI High Level Expert Group and an AI Watch to help Europe become the world-leading region for developing and deploying cutting-edge, ethical and secure AI.

The emphasis put on AI over the past year has stressed the importance of the digital economy and, indeed, of the data economy as no real AI could truly work without data. The centrality of the Data Economy as a key objective of European policies and, specifically, of the Digital Single Market (DSM) Strategy, is therefore bound to remain unchallenged but new issues and potentially impactful changes linger on the horizon. In 2019 both the European Commission and the European Parliament will end their mandate. This is likely to introduce new policy priorities but, the need for Europe to capture the digital opportunity will remain at the forefront of European strategies and crucial decisions that will influence the way digital technologies will be developed and deployed will be taken.

These choices, as well as the influence of external factors, may further push or hinder the growth potential of the European Data Market and the European Data Economy over the next few years. Our scenarios present the possible, alternative evolution paths driven by different macroeconomic and framework conditions and shaped by critical turning points to be faced in the next years by governments, businesses and social actors. The uncertainty surrounding Brexit, however, will persist in 2019 as every option remains on the table at the time of writing.

From an international perspective, Europe continues to lag behind the U.S. in terms of the size of the Data Market and the Data Economy, but it displays a remarkable potential that will have to be carefully nurtured over the next few years. In this respect, Europe's efforts in setting standards not only as far as ethics principles are concerned, but also with regards to data protection, privacy and fair competition, are gaining international attention and can set the grounds for Europe to be a role model for other countries. Europe's firm and pioneering position supporting individuals' control over their own data and profits coming from them, for instance, could prompt the set-up of sharable standards which, if adopted on a wide-scale, could benefit users and constrain the power of the big tech corporations (among which the above mentioned GAFA) currently exploiting the power of data in what is a semi-monopolistic way.

METHODOLOGICAL ANNEX

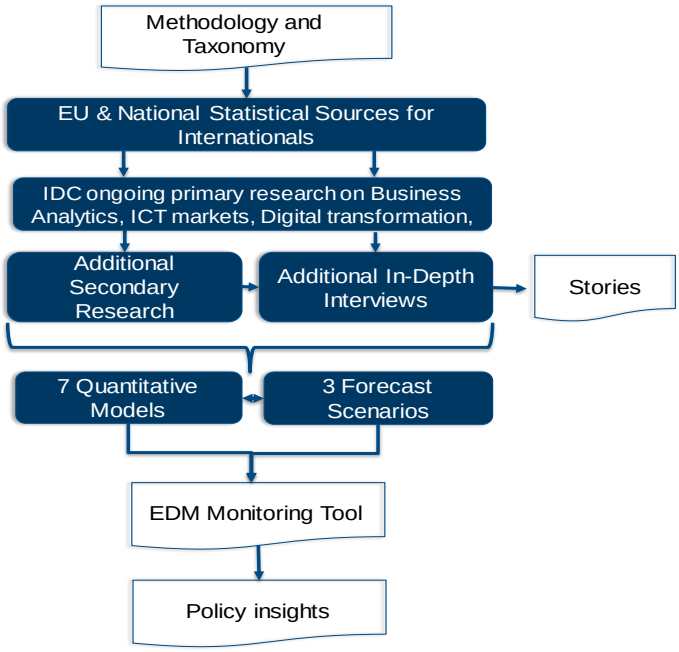
Overview

In line with the methodology adopted in the previous European Data Market Study (SMART 2013/0063), the measurement methodology for this updated report was based on the steps outlined in Figure 15 below. Compared to the previous steps it does not include the ad-hoc surveys which were used to establish the baseline. However, thanks to the use of IDC primary research data tracking the market, we have already proven the feasibility of updating the indicators without repeating the initial surveys.

The main steps of the methodology did include:

- Desk research on the main EU and global national and statistical sources; each indicator has specific set of sources;
- Extraction of data from the relevant IDC surveys and databases;
- Additional secondary research and case studies interviews for the stories, which in turn did feed back to the indicator models to help in the modelling and estimate of indicators;
- A selected number of opinion leader and stakeholder interviews to feed into the modelling and scenario assumptions;
- Implementation of the 7 indicators models and elaboration of results;
- Development of the forecast scenario assumptions and update of the 3 scenarios;
- Assessment of policy insights building on the results of the previous steps.

Figure 17: A sophisticated Methodology



Desk Research

As done in the first study, the study team reviewed the list of relevant public sources and updated it to collect additional relevant data. The list of the main sources used so far is outlined below and will be constantly revised when preparing the second and the third round of measurement of the indicators.

- Concerning the indicators on Data Market, data companies, data companies' revenues, and the Data Economy the main sources were:
    - Eurostat business demography statistics in the European Union, treating aspects such as the total number of active enterprises in the business economy, their birth rates, death rates, and the survival rate (last update: December 2014);
    - Eurostat annual structural business statistics with a breakdown by size-class are the main source of data for an analysis of SMEs (latest update: March 2016);
    - IDC's detailed market forecast estimates for IT Hardware, Software, and IT Services from 2014 and 2015;
    - IDC Worldwide Black Book (Standard Edition), quarterly updates form the years 2014 through 2015. The Black Book represents IDC's quarterly analysis of the status and projected growth of the worldwide ICT industry in 54 countries.
    - IDC End-User IT Trends and Digital Transformation: IDC European Vertical Markets Survey 2015
    - IDC European Vertical Markets Survey, 2014: More Western European SMBs Will Invest in Software Solutions Beyond Maintenance, July 2015
    - IMF World Economic Outlook (WEO) Database, April 2016
    - Consensus Forecasts, Consensus Economics, monthly updates, September 2015 – March 2016.
- For the data professionals we used in addition the following sources:
    - OECD Digital Economy Papers, among which: OECD (2014), Measuring the Digital Economy: A New Perspective; OECD Publishing.
    - ILOSTAT (International Labour Organization) Statistics and Databases (2015)
    - EUROSTAT Tertiary Education Statistics (Last update: December 2015).
    - European Data Science Academy (EDSA) project deliverables and publications (July 2015).
- For the indicator on the Citizens' Reliance on the Data Market we used in addition the following sources:
    - The Digital Economy and Society Index (DESI), Human Capital Dimension, (2a Basic Skills and Usage; 2b Advanced skills and Development), last update, February 2016.
    - IDC European Quarterly Wearables Tracker Results: Western Europe 3Q15 Analysis, January 2016
    - IDC FutureScape: Worldwide Wearables 2016 Predictions, November 2015.

Ad Hoc Workshop and Expert Interviews

For the update of the methodology and of the assumptions for the indicators models and the forecast scenarios, the study team has carried out a few selected, one-to-one interviews with key experts and has organised a specific workshop in collaboration with the BDVe project.

The workshop on "The European Data Economy by 2025" was led by IDC on October 20th, at the BDVA's premises in Brussels. It gathered insights from the high-level group of industry and research experts, from the BDVA community, about the potential growth paths of the European Data Economy by 2025. The group identified and discussed its possible key turning points, and the most impactful drivers and barriers of the data-driven innovation.

Four thematic areas were discussed in the plenary session and in ad-hoc breakout sessions:

- Supply and demand dynamics;
- Technology trends;

- Policy and regulation;
- Social role of data.

A list of 21 different evolution paths of the data economy was identified and a narrative developed for each of these possible paths. A subsequent discussion in the afternoon successfully trimmed down the list to a manageable number of seven fully developed scenarios on the basis of which, the assumptions underpinning the different set of indicators were updated.

In the course of the workshop, the study team also conducted a few selected interviews with key experts to:

For the update of the methodology and of the assumptions for the indicator models and the forecast scenarios, we carried out 4-5 key expert interviews in the first phase of the study focused on:

- Validation of the methodology approach
- Feedback on main factors affecting the dynamics of data-driven innovation
- Suggestions of improvement

Interviews were carried out with:

- Paul Czech, Know-Center GmbH, Research Center for Data-Driven Business & Big Data Analytics
- Anthoine Dusselier, Dawex
- Bas Kotterink, TNO
- Philip Carnelley, IDC

The initial list of experts identified in the proposal phase of the project, and already leveraged during the preparation of the previous study, remains valid and will be used in the upcoming rounds of measurements of the indicators. The list includes:

- Vincenzo Spiezia, Head of the Information and Communication Technologies Unit in the Directorate for Science, Technology and Industry of the OECD
- Jonathan Cave, Senior teaching fellow in economics, University of Warwick
- Elena Simperl, professor of Computer Science at the University of Southampton and manager of the European Data Science Academy (EDSA) project

Additional Research on Start-ups and DIH

In the framework of this updated study, IDC has conducted extensive and in-depth desk research on the diffusion of data-driven start-ups and different types of already existing supporting initiatives and entities by relying on a valuable number of primary and secondary resources including, but not limited to:

- Business incubators, accelerators, DIHs own websites and directories;
- European and National Statistic's Offices' sources such as Eurostat's Business demography statistics and other statistics from national offices presenting data on business demography and treating aspects such as the total number of active enterprises in the business economy, their birth rates, death rates, and the survival rate;
- Interactive maps such as Start-ups Hub Europe (http://www.startuphubs.eu/), which provides an updated and comprehensive mapping of Europe's start-up ecosystem, and the Digital Innovation Hubs Tool (http://s3platform.jrc.ec.europa.eu/digital-innovation-hubs-tool);

- Additional sources from the European Commission such as the Entrepreneurship 2020 Action Plan and the Start Up Europe initiative under the Digital Single Market strategy (https://ec.europa.eu/digital-single-market/en/policies/startup-europe);
- Other sources from the specialized press or the business community such as EU-Startups.com (one of the leading start-up blogs in Europe);
- Ad-hoc studies such as the First and the Second European Startup Monitor carried out by the German Startups Association (GSA) and the European Startup Network (ESN).

For the sake of clarity and completeness, we provide the following definitions of the main types of supporting initiatives and actors identified during the research activity:

Key Terms - Definitions

Business accelerator A program or organisation characterised by the following common traits: i) an application process that is open to all, yet highly competitive; ii) possible provision of pre-seed investment (grant or equity); iii) a focus on small teams instead of individual founders; iv) time-limited support (usually from 3 to 6 months) comprising programmed events and intensive mentoring; v) cohorts or 'classes' of start-ups rather than individual companies.

Business angel A private individual, often of high net worth, and usually with business experience, who directly invests part of his or her personal assets in new and growing private businesses. Business angels can invest individually or as part of a syndicate where one angel typically takes the lead role.

Besides capital, angel investors provide business management experience, skills, and contacts for the entrepreneur. Experienced angels also know that they may have to wait for a return on their investment. They can therefore be a good source of 'smart and patient' capital.

In many countries, they constitute the second largest source of external funding in newly established ventures, after family and friends.

Business incubator An organisation designed to advance the growth of start-ups entering the beginning stages of building their company through an array of business support resources and services that could include physical space, coaching, common services, and networking connections. Incubators operate on an open-ended timeline (usually from 1 to 5 years): they focus more on the longevity of a start-up and are less concerned with how quickly the company grows. Incubators do not traditionally provide capital to start-ups and are often funded by universities or economic development organizations. They also don't usually take an equity stake in the companies they support.

Digital Innovation Hubs (DIHs) Ecosystems that consist of SMEs, large industries, start-ups, researchers, accelerators, and investors, which aim to create the best conditions for long-term business success. DIHs help companies to become more competitive in terms of their business/production processes, products or services using digital technologies. They are based upon technology infrastructure and provide access to the latest knowledge, expertise and technology to support their customers with piloting, testing and experimenting with digital innovations. DIHs also provide business and financing support to implement these innovations, if needed across the value chain. A DIH is a regional multi-partner cooperation (including organisations like RTOs, universities, industry associations, chambers of commerce, incubator/accelerators, regional development

agencies and even governments) and can also have strong linkages with service providers outside of their region supporting companies with access to their services.

Venture capital Innovative and growth-oriented small businesses need to raise capital (equity investment) from external sources because they do not have their own resources or cannot access loans. Firms typically look for venture capital to provide them with the financing they need to expand, break into new markets, and grow faster. Although venture capital is only relevant for a small group of firms, it is essential for the growth of innovative firms. Venture capital funds raise a large part of their funding from institutional investors and they usually invest large amounts into firms with the potential for rapid growth.

Forecast Scenarios

In our methodology, the scenarios are used to elaborate the potential alternative growth paths of the European Data Economy, taking into account the main economic, technological, policy-regulatory and social factors affecting its development. The qualitative scenarios interact with the quantitative forecast models with a mutual fine-tuning and validation effect, by investigating the rationale behind potential growth trends, and vice-versa by taking into account insights from the data. The ultimate objective of the scenarios, however, is to analyse which combination of framework conditions and policy actions may maximise the growth potential of the European Data Market and Economy, and by feeding into the models estimate the actual size and depth of the potential benefits. In this way the scenarios provide a realistic approach to the forecast estimates – since we project a range of values (not a single estimate which may be widely off the mark) - and provide guidance on the potential consequences of different external events or alternative policy choices.

We have implemented the same scenario methodology used in the previous EDM Study. The scenario model used in this study is based on the definition of alternative assumptions about four main groups of key factors. IDC has developed and implemented this model in several projects about various ICT markets, from the Future Internet to Cloud Computing and the IoT and we believe it is thoroughly validated.

Every year the assumptions within each of the main groups of factors will be revised and updated or validated, or new ones will be added, leveraging the results of the research as well as IDC's periodically updated global Market Forecast Assumptions. These assumptions are collected, aggregated and shared with all IDC analysts at a global level by IDC's Global Research Group which is composed of experienced analysts and economists.

The selection of the most relevant factors in the scenario model was based on two main criteria:

- High level of impact on the development of the Data Market;
- High level of uncertainty, with potential different outcomes (assumptions) over the next five years.

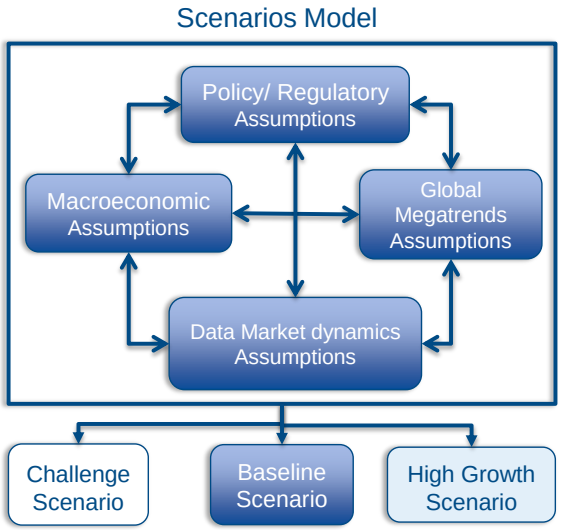
The four main groups of factors are:

- Macroeconomic factors;
- Policy/regulatory conditions;
- Data Market dynamics factors;
- Global megatrends affecting all technology markets.

Even though they may seem obvious, still these four clusters correspond to the main typologies of factors which affect the evolution of the Data Market. Each cluster aggregates a set of interrelated

key factors; their combination differentiates the three scenarios (Figure 16). The scenarios are characterised by the interaction and co-dependency of these factors; no scenario can be explained only by one factor or one group of factors, not even GDP growth.

Figure 18: Structure of the Scenarios Model



Source: European Data Market Monitoring Tool, IDC 2015

The scenario model and the forecast indicators models are correlated.

The table below summarises the rationale of their selection and how their assumptions were used as inputs to the indicators' forecast models.

Table 18: Identification of main Factors driving the Scenarios

Key Factors	Rationale	Inputs to the Forecast Models
Macroeconomic factors	Strong influence of the macroeconomic context on Data Market growth	Alternative forecasts of: EU GDP growth 2016-2025 ICT spending growth 2016-2025 Alternative economic growth conditions
Policy/Regulatory conditions	Strong influence of the policy/regulatory framework on the model of development of the Data Market	Alternative policy and regulatory conditions by scenario
ICT Market/ FIWARE dynamics factors	Strong influence of alternative supply-demand dynamics on the market development paths	Alternative supply and take-up models by scenario
Global megatrends	Strong influence of global digital innovation trends on the EU Data Market growth	Alternative assumptions on the development of IoT, Cloud Computing, Mobile technologies based on IDC's 2025 forecasts

Measuring Data Professionals

Definition and Scope

Data professionals *are workers who collect, store, manage, and/or analyse, interpret, and visualise data as their primary or as a relevant part of their activity. Data professionals must be proficient with the use of structured and unstructured data, should be able to work with a huge amount of data and be familiar with emerging database technologies.*

In our definition, data professionals are not only data technicians but also data users who, based on more or less sophisticated tools, take decisions about their business or activity, after having analysed and interpreted available data. According to our definition, data professionals belong to the category of knowledge workers and specifically "codified" knowledge workers (Lundavall and Johnson, 1994); data professionals specifically deal with data while knowledge workers deal with information and knowledge.

The indicator has been measured according to the segmentations presented in the following table, including two sub-indicators about the share on employment and the intensity of employment.

Table 19: Indicator 1 – Data Professionals

Indicator 1 – Data Professionals				
N.	Name	Description	Type and Time	Segmentation
1.1	Number of data professionals	Total number of data professionals in the EU	Number, 2016-17-20 Forecast to 2025, 3 Scenarios	By Geography: 28 EU MS + total EU
				By Industry: 11 industry sectors NACE rev.2
1.2	Employment share	Total number as a share of total employment in the EU	% of total employment, 2017-18-19	By Geography: 28 EU MS + total EU
				By Industry: 11 industry sectors NACE rev.2
				By Size: not applicable
1.3	Intensity share	Average number of data professionals per company (only for private sector)	Number, 2017-18-19	By Geography: 28 EU MS + total EU
				By Industry: 11 industry sectors NACE rev.2
				By Size: not applicable

Methodology Approach

Our approach is based on an iterative process and on a calibration process of the final estimates. The approach has been repeated in the new study based on updates of the main sources.

Statistical Identification

Data professionals are not classified as such into any of the labour and occupation statistics. In order to define them statistically, we have adopted the International Standard Classification of Occupations (ISCO-08), selecting categories where data professionals may be included. The criteria adopted for the selection of the ISCO-08 codes are the following:

- We have selected the occupations where data professionals can be involved either as data providers or as data users;
- We have selected the occupations from 1 to 4-digit disaggregation;
- The occupation codes selected are those where the presence of data professionals can be detected because:
    - They hold deep analytical skills;

- They do not need deep analytical skills but basics understanding of statistics and/or machine learning in order to conceptualise the questions that can be addressed through deep analytical skills;
- They are the ones providing enabling technology and therefore they are providers of data services.

- The selected codes are those where a significant part of the workers may be data professionals; the occupations where the data professionals are a very marginal part of the workers have been excluded; as an example, the medical practitioners have been excluded, although some practitioners may be data professionals because they undertake research activities. Since they are only a very marginal part of the practitioners, we excluded them from the occupations where data professionals are present;
- We excluded all the data professionals which are not included into the knowledge economy perimeter because their occupation is a low skilled one, i.e. with high routine level (as an example, call centre workers are in theory data professionals but since their activity is a routine one and as such excluded from the knowledge economy, they are not considered data professionals).

Table 20: ISCO-08 Structure and Data Professionals

	ISCO-08 structured Classification			
	Major Groups (1 digit)	Sub-groups (2 digits)	Minor Groups (3 digits)	Units (4 digits)
Number of codes ISCO-08 structure	10	43	130	436
Number of selected codes including data professionals	4	9	21	52
Share of data professionals' codes in the ISCO-08 structure	40%	21%	16%	12%

Source: IDC elaboration on ISCO codes

Calculation of the quantitative Perimeter

The quantitative perimeter of employment where data professionals are trackable is based on the selected ISCO codes crossed with the NACE classification of economic activities, for each one of the 28 Member States and the EU as a whole, and has been updated based on the sources updates.

Estimate and Calibration of the Penetration of Data Professionals

The next step is the estimate of percentage of data professionals within the perimeter of data professional candidates. To calculate the coefficients for the calculation of such %, we have elaborated a set of assumptions (specified in the D2- Methodology report of the EDM Study). The assumptions have been revised and updated for each release of the study and applied to the model to calculate the share of data professionals by Member State and by industry.

Forecasting Data Professionals

The same model was applied to forecast data professionals to 2025, by developing specific assumptions by scenario, even though the level of uncertainty is higher, and the reliability of the forecasts is lower.

Measuring Data Companies

Definition and Scope

Data companies *are organisations that are directly involved in the production, delivery and/or usage of data in the form of digital products, services and technologies. They can be both data suppliers' and data users' organisations:*

- ***Data suppliers*** *have as their main activity the production and delivery of digital data-related products, services, and technologies. They represent the supply side of the Data Market.*
- ***Data users*** *are organisations that generate, exploit collect and analyse digital data intensively and use what they learn to improve their business. They represent the demand side of the Data Market.*

Table 21: Indicator 2 – Number of Data Companies

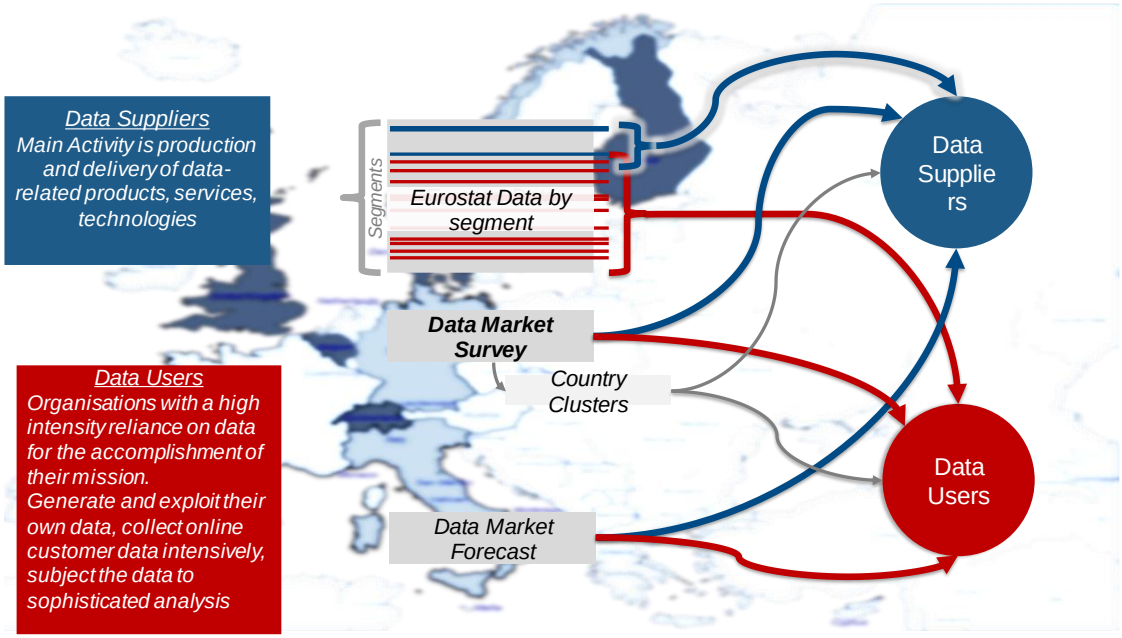
Indicator 2 – Data companies				
N.	Name	Description	Type and Time	Segmentation
2.1	Number of data suppliers	Total number of data suppliers, measured as legal entities based in the EU	Number, 2017-18-19 Forecast to 2025, 3 Scenarios	By Geography: 28 EU MS + total EU
				By Industry: 2 NACE rev2 Section J Information and Communication and section M Professional, scientific and technical activities
				By Company Size: below 250 employees above 250 employees
2.2	Share of data suppliers	Total data companies on total companies in industry J and M	% 2017-18-19	By Geography: 28 EU MS + total EU
				By Industry: 2 NACE rev2 Section J Information and Communication and section M Professional, scientific and technical activities
2.3	Number of data users	Total number of data users in the EU, measured as legal entities based in one EU country	Number, 2017-18-19 Forecast to 2025, 3 Scenarios	By Geography: 28 EU MS + total EU
				By Industry: 11 industry sectors NACE rev.2
				By Company Size: below 250 employees above 250 employees
2.4	Share of data users	Total data users as share of total private companies	% 2017-18-19	By Industry: 11 industry sectors NACE rev.2

Methodology Approach

Data companies have been measured by updating the same model used in the previous EDM Study (see Figure below) which leverages both IDC and public sources.

- Eurostat business demography statistics in the European Union, treating aspects such as the total number of active enterprises in the business economy, their birth rates, death rates, and the survival rate (last update: December 2014);
- Eurostat annual structural business statistics with a breakdown by size-class are the main source of data for an analysis of SMEs (latest update: March 2016);
- IDC's detailed market forecast estimates for IT Hardware, Software, and IT Services from 2014 and 2015;
- IDC Worldwide Black Book (Standard Edition), quarterly updates form the years 2014 through 2015. The Black Book represents IDC's quarterly analysis of the status and projected growth of the worldwide ICT industry in 54 countries.
- IDC End-User IT Trends and Digital Transformation: IDC European Vertical Markets Survey 2015
- IDC European Vertical Markets Survey, 2014: More Western European SMBs Will Invest in Software Solutions Beyond Maintenance, July 2015

Figure 19: Data Companies Model



Measuring the Revenues of Data Companies

Definition and Scope

Data companies' revenues *are the revenues generated by data suppliers for the products and services specified in our definition of the Data Market. The revenues correspond to the aggregated value of all the data-related products and services generated by Europe-based suppliers, including exports outside the EU.*

Table 22: Indicator 3 – Revenues of Data Companies

Indicator 3 – Revenues of Data Companies

N.	Name	Description	Type and Time	Segmentation
3.1	Total revenues of data companies	Total revenues of the Data Suppliers calculated by Indicator 2	Billion €, 2017-18-19 Forecast to 2025, 3 Scenarios	By Geography: 28 EU MS + total EU By Company Size: below 250 employees above 250 employees
3.2	Share of data companies' revenues	Total revenues of the Data Suppliers calculated by Indicator 2	% of revenues on total, 2017-18-19	By Geography: 28 EU MS + total EU

Methodology Approach

The indicator has been measured applying the same model used in the previous EDM Study, which calculated the revenues by feeding on:

- Eurostat and IDC statistics on average IT vendors revenues by size and sector;
- The total number of data companies by country, industry and size class;
- The value of the Data Market by country and industry;
- The estimated share of exports-imports in the value of the Data Market.

Measuring the Data Market

Definition and Scope

The Data Market is the marketplace where digital data is exchanged as "products" or "services" as a result of the elaboration of raw data. We define its value as the aggregate value of the demand of digital data without measuring the direct, indirect and induced impacts of data in the economy as a whole. The value of the Data Market includes imports (data products and services bought on the global digital market from suppliers not based in Europe) and excludes the exports of the European data companies.

Table 23: Indicator 4 – Size of the Data Market

Indicator 4 – Size of the Data Market				
N.	Name	Description	Type and Time	Segmentation
4	Value of the Data Market	Estimate of the overall value of the Data Market	Billion €, 2017-18-19 Forecast to 2025, 3 Scenarios	By Geography: total EU, EU28 By Industry: 11 industry sectors NACE rev.2 By Size: not applicable

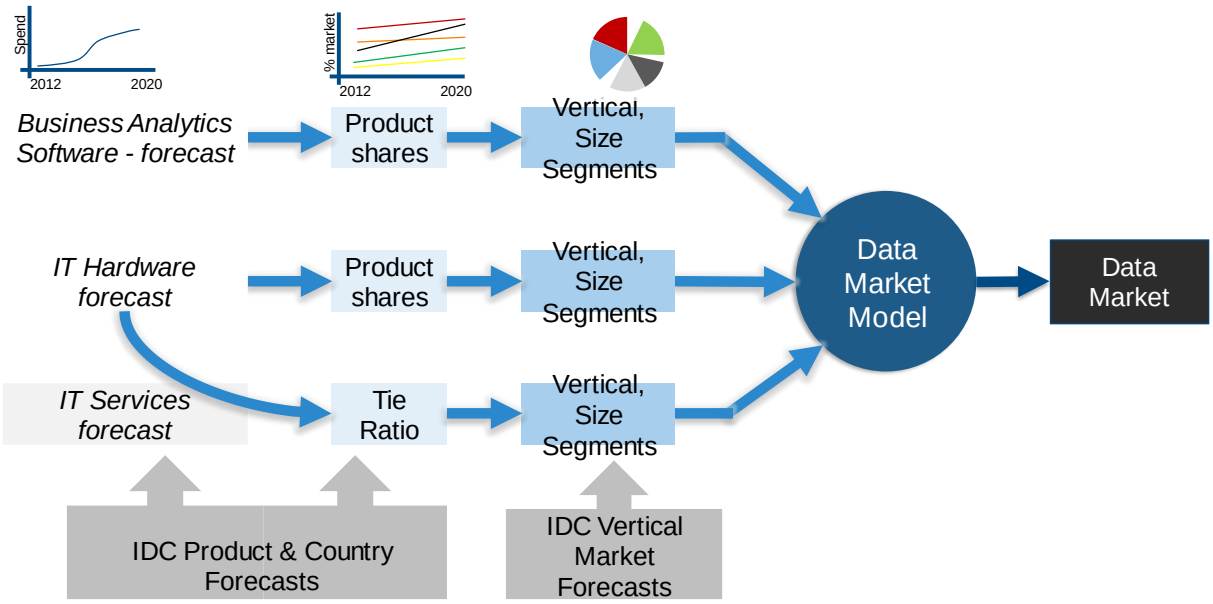
Methodology Approach

The Data Market indicator is being updated every year for the duration of the study. The model is based on the extraction of data from IDC databases concerning the components of hardware,

software and services spending which fall in the definition of the Data Market. The IDC data is already segmented by country and by industry, even though not all Member States are covered, and the industry classification is slightly different from the one proposed in this project. The respective shares for the software, hardware, and services market used to derive the Data Market are derived from IDC surveys covering Big Data, IT spending patterns and intentions in the European market, and a survey of data suppliers and data users in key Member States, together with analyst expertise and alignment with IDC's European and worldwide forecasts for the business analytics and Big Data Market.

The model updates the Data Market value shares by Member State and by industry.

Figure 20: Data Market Model



Source: IDC 2016

Measuring the Data Economy

Definition and Scope

*The **Data Economy** measures the overall impacts of the Data Market on the economy as a whole. It involves the generation, collection, storage, processing, distribution, analysis elaboration, delivery, and exploitation of data enabled by digital technologies. The Data Economy also includes the direct, indirect, and induced effects of the Data Market on the economy.*

The Data Economy indicator measures the value of the Data Economy based on the estimate of all the economic impacts following the adoption of data-driven innovation and data technologies in the EU. As such, the indicator aggregates direct, indirect, induced impacts of the Data Market defined as follows.

1. **The direct impacts**: these are impacts generated by the data industry itself; they represent the activity engendered by all businesses active in the data production. **The quantitative direct impacts are measured by the revenues from data products and services sold, i.e. the value of**

the Data Market. We prefer to adopt the Data Market value as a good proxy of the direct impacts because its estimates are more reliable than the value of the revenues.
The direct impacts: the initial and immediate effects generated by the data suppliers; they represent the activity potentially engendered by all businesses active in the data production. **The quantitative direct impacts have then been measured as the revenues from data products and services sold, i.e. the value of the Data Market.** As Data Market estimation is more reliable than data companies' revenues estimation, we consider the Data Market value as a good proxy of the direct impacts. Therefore, for the sake of simplicity, direct impacts do coincide with the value of the Data Market.

2. **The indirect impacts:** the economic activities generated along the company's supply chain by the data suppliers. There are two different types of indirect impacts: the backward indirect impacts and the forward indirect impacts (Richardson, 1985):
    a. **The backward indirect impacts**: such impacts represent the business growth resulting from changes in sales from suppliers to the data industry. In order to produce and deliver data products and services, the data companies need inputs from other stakeholders. Revenues from those sales to data companies are the backward indirect impacts.
    b. **The forward indirect impacts**: such impacts include the economic growth generated through the use of data products and services by the downstream industries, i.e. the data users as a selected number of industries. For the user companies, data is now a relevant factor of production; the adoption of data products and services by the downstream industries provides different types of competitive advantage and productivity gains to the user industries. The main benefits that the exploitation of data can provide to downstream industries are (OECD, 2013, Mc Kinsey, 2011):
        i. Optimising production and delivery processes: data-driven processes (data-driven production);
        ii. Improving marketing by providing targeted advertisements and personalised marketing practices (data-driven marketing);
        iii. Improving existing organisation and management practices (data-driven organisation).
3. **The induced impacts**: these impacts include the economic activity generated in the whole economy as a secondary effect. Induced additional spending is generated both by new workers, who receive a new wage, and by the increased wage of existing jobs. This spending induces new revenues creation in nearly all sectors of the economy. The additional consumption does support economic activity in various industries such as retail, consumer goods, banks, entertainment, etc.

Table 24: Indicator 5 – Value of the Data Economy

Indicator 5 – Value of the Data Economy				
N.	Name	Description	Type and Time	Segmentation
5	Value of the Data Economy	Value of the Data Market plus direct, indirect and induced impacts on the EU economy	Billion €, 2017-18-19 Forecast to 2025, 3 Scenarios	By Geography: Total EU + EU 28
5.1	Incidence of the Data Economy on GDP	Ratio between value of the Data Economy and EU GDP	%, 2017-18-19 Forecast to 2025, 3 Scenarios	By Geography: EU 28 + Total EU

This estimate of the Data Economy does not include the user benefits and social impacts of data-driven innovation such as changes in quality of life (health, safety, recreation, air quality). Although these benefits may be evaluated in economic (money) terms, they are not economic impacts as such and as defined above as they do not induce an increase in the business activities and a consequent growth in GDP.

Analysts underlined that the new decision-making processes act as a rationalisation and optimisation factor (Brynjolfsson, 2011, Mc Kinsey, 2012), since they improve effectiveness and efficiency, and in some cases, they may have a disruptive effect. The impacts related to the new decision-making processes are the one we have called the forward indirect impacts.

The value creation process based on data rests on the elaboration of information and knowledge (OECD 2016), although the boundaries between data, information, and knowledge are sometimes fuzzy. The huge volume of data is a global phenomenon which is sometimes viewed with suspicion by citizens, consumers and businesses because data flows are seen as an intrusion of the privacy. Nevertheless, there is currently some evidence showing that data analysis can provide benefits to both businesses and consumers. By the way, this is not surprising since we should remind that the economic theory holds that information encourages competition between businesses for the benefit of consumers.

Data do not provide value and benefits as such; data need to be collected, stored, aggregated, combined and analysed in order to be appropriately used for decision making processes. To create value, data need to be processed (OECD, 2016):

- **Extracting information from structured and unstructured data**: data analytics techniques are today able to analyse both structured and unstructured data. We should remind here that most data stored by businesses are unstructured (IDC, 2012). Technologies such as optical character recognition, natural language processing, face recognition algorithms and machine learning algorithms are empowering the use of all data.
- **Real-time monitoring and tracking**: analysis of data in real time is often mentioned as one of the most powerful factor since it supports organisations to make real-time decisions, which, in a fast-changing world, is a well-known competitive advantage.
- **Inference and prediction**: until now, prediction was based exclusively on prior information and data series. Data analytics can now enable the creation of information even without prior information. Such information can be created through patterns and correlations of data. Personal information, for example, can be deduced from anonymous or non-personal data. Businesses and organisations demand real time insights rather than historical and periodical information, and for advanced specialised data analytics services. Algorithms allow machine and statistical learning based on non-specific data; businesses can learn and predict a lot about their customers even if they do not have specific data and time series about the issue they are interested in. Machine learning has, as an example, applications in health care where data collected on patients are recorded by imaging, or it supports production processes to increase the quality of production.

The diffusion of technology supporting production and analysis of data induces organisations and businesses to base their decisions on data much more than they were used to do. As pointed out by OECD in its recent report, the process to take decisions is also changing. Decision makers do not necessarily need to understand the phenomenon before they act on it. A store can change the product placement based on data analysis without the need to know the reason why such a change should improve the sales. There is therefore a decision automation process: “first comes the analytical factor, then the action, and last, if at all, the understanding” (OECD, 2015).

The impacts of such a new approach to decision making and to the use of data in all the enterprises and organisations' functions are many and varied, so that we believe, such impacts will be object of studies and analysis in the upcoming years. It is, at this point, difficult to classify them and to suggest a taxonomy of such impacts.

Such impacts have been observed through some empirical studies and case analysis. The most relevant ways the benefits appear are the following.

- **Creating more information, knowledge and transparency**: technology is making data more accessible and exploitable to all kind of stakeholders, including SMEs. This increases transparency and decisions are made on a rational process.
- **Improving performance**: having access to a wide information and to a high number of data is changing the way of making decisions. An increasing number of organisations are going to become data-driven organisations, which means that they make decisions based on empirical results. As an example, retailers can adjust prices and promotions, more precisely than they were used to and in real time. This may improve competitiveness. McKinsey underlines that the health sector is achieving a lot of benefits from the new making decisions process: studies on clinical data allow to identify and understand the sources of variability in treatment, to identify the best treatment protocols and to create guidelines for the optimization of treatment decisions. This does not only increase the effectiveness of treatments, but it also produces saves.
- **Improving customization of actions for better decisions**: data technology is definitely improving the segmentation of customers and the analysis of their preferences in real time. This allows companies to supply products and services targeted to specific groups of individuals who have specific needs and preferences. Such a segmentation is also useful when supplying public services. Such a segmentation helps define the price precisely and offering exactly what is needed which means a better quality and also companies avoid offering products and services the consumers are not willing to pay.
- **Innovating products and services as well as business models**: the more information and understanding businesses have about their customers, the better they can serve them. It is important to say that although consumers may fear their privacy is injured, this can also provide them unexpected surplus: real time price comparison services do not only provide better transparency but also allow buying the best product at the most convenient price (for example when buying online airline tickets or when booking hotels). Companies can in fact produce and create new products and services to better satisfy their customers' needs. This is true also for the public sector and specifically for the health care system where preventing care programs can be created.

These effects are reflected in an increase in revenues due to higher market share from the increase in competitiveness or due to a reduction in costs. All these effects are included in the forward indirect impacts; these impacts are delivered on the user industry, and because of the above reasons, these are the impacts we consider new on the overall economic system.

Methodology Approach

Measuring the Data Economy depends on the macroeconomic context on one hand, and on the adoption/diffusion and integration processes the companies are implementing on the other hand. Moreover, there is a necessary time lag before the impacts take place in the economic system. Therefore, the estimates are based on a set of assumptions, including choices about proxy indicators.

In order to measure the impact of the diffusion and use of data services and products, we estimated each component (as defined in the above paragraph) of the impact separately.

The estimation approach developed in the previous study was based on a number of assumptions on one hand and on results from a survey launched during the first-year research.

The following assumptions have been confirmed:

- The penetration rates of data in terms of value added for the user industries using data are positively correlated to the penetration rate in terms of number of companies using data.
- The survey conducted in the study 2013-2016 provided information about the quantitative benefits due to the use of data, for the six major Member States plus Czech Republic; such benefits have been taken into consideration for the six major Member States.
- For Austria, Belgium, Denmark, Finland, Ireland, Luxembourg, Malta, the Netherlands, and Sweden we assumed that these Member States have the same distribution of benefits as the average of the Big Six.
- For the other Member States, we estimated the benefits of the rest of Europe, based on the survey results, and we assumed that all the minor Member States are achieving benefits similar to the rest of Europe.
- For the induced impacts, we assumed that the additional earnings are spent according to the general economic mood.

In order to update the estimates of the different components of the impacts, we have adopted some new assumptions:

- In the next three years, we are going to stay in a relatively emerging stage of the data diffusion, so that in our view the structure of the data impacts is not going to change.
- For the quantitative benefits due to the use of data, we assume that the benefits will quantitatively vary and be correlated to the macroeconomics trends and specifically with the industries' trends (and stakeholders) affected.

Measuring the Data Skills Gap

Definition and Scope

This indicator captures the potential gap between demand and supply of data skills in Europe, since the lack of skills may become a barrier to the development of the data industry and the rapid adoption of data-driven innovation.

Table 25: Indicator 6 – Data Skills Gap

Indicator 6 – Data Skills Gap				
N.	Name	Description	Type and Time	Segmentation
6	Data Workers Skills Gap	Gap between demand and supply of data workers	Absolute number and % on total demand, 2017-18-19 Forecast to 2025, 3 scenarios	By Geography: total EU28; main EU Member States

Methodology Approach

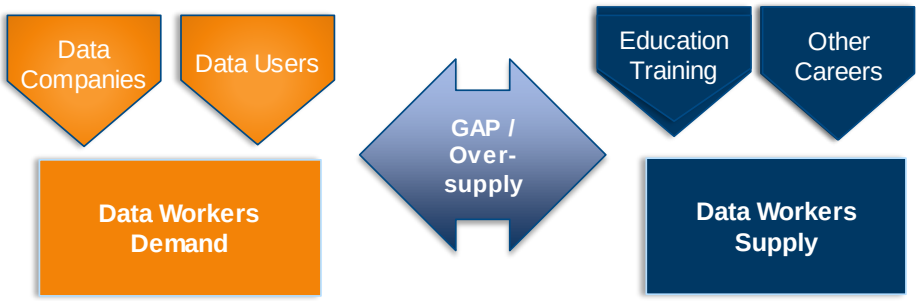
The methodology approach is the same implemented by IDC-empirica to estimate the supply-demand balance of ICT skills in the EU (e-Skills) on behalf of the EC DG Enterprise (now DG GROW). The model was first developed in 2009 and since then has been successfully validated and updated

several times. The results have been used by the EC to support the e-skills policy and the latest results were presented in December 2015 at the European E-skills 2015 Conference in Brussels[6]. However, data skills are not a subset of ICT skills so the scope of supply and the dynamics of demand are different from the e-skills model developed by IDC.

To update the measurement of the indicators the study team has applied the same model developed for the previous EDM Study, combining the estimates and forecasts of the demand and supply of data professionals with data skills leveraging a wealth of different sources, among which:

- OECD Digital Economy Papers, among which: OECD (2014), Measuring the Digital Economy: A New Perspective; OECD Publishing.
- ILOSTAT (International Labour Organization) Statistics and Databases (2015)
- EUROSTAT Tertiary Education Statistics (Last update: December 2015).
- European Data Science Academy (EDSA) project deliverables and publications (July 2015).

Figure 21: The Data Skills Demand-Supply Balance Model



Source: European Data Market Monitoring Tool, IDC 2016

[6] "e-Skills in Europe: Trends and Forecasts for the European ICT Professional and Digital Leadership Labour Markets (2015-2020)", empirica Working Paper (November 2015)

ESSENTIAL GLOSSARY – THE KEY INDICATORS

Data professionals *are data workers who collect, store, manage, and/or analyse, interpret, and visualise data as their primary or as a relevant part of their activity. Data professionals must be proficient with the use of structured and unstructured data, should be able to work with a huge amount of data and be familiar with emerging database technologies. They elaborate and visualise structured and unstructured data to support analysis and decision-making processes.*

Data companies *can be both data suppliers' and data users' organisations:*

- ***Data suppliers*** *have as their main activity the production and delivery of digital data-related products, services, and technologies. They represent the supply side of the Data Market.*
- ***Data users*** *are organisations that generate, exploit, collect and analyse digital data intensively and use what they learn to improve their business. They represent the demand side of the Data Market.*

Data companies' revenues *are the revenues generated by data suppliers for the products and services specified in our definition of the Data Market. The revenues correspond to the aggregated value of all the data-related products and services generated by Europe-based suppliers, including exports outside the EU.*

The ***Data Market*** *is the marketplace where digital data is exchanged as "products" or "services" as a result of the elaboration of raw data. We define its value as the aggregate value of the demand of digital data without measuring the direct, indirect and induced impacts of data in the economy. The value of the Data Market includes imports (data products and services bought on the global digital market from suppliers not based in Europe) and excludes the exports of the European data companies.*

The ***Data Economy*** *measures the overall impacts of the Data Market on the economy. It involves the generation, collection, storage, processing, distribution, analysis elaboration, delivery, and exploitation of data enabled by digital technologies. The Data Economy also includes the direct, indirect, and induced effects of the Data Market on the economy.*

The ***Data Professionals' Skills Gap*** *captures the potential gap between demand and supply of data skills in Europe, since the lack of skills may become a barrier to the development of the data industry and the rapid adoption of data-driven innovation.*

Data *is usually defined as qualitative or quantitative statements or information which can be coded and which are assumed to be factual and not the product of analysis or interpretation. For the sake of this study we consider only data which is collected, processed, stored, and transmitted over digital information infrastructures and/or elaborated with digital technologies. This definition includes multimedia objects which are collected, stored, processed, elaborated and delivered for exploitation through digital technologies (for example, images databases).*

Information *is the output of processes that summarise, interpret or otherwise represent the content of a message to convey meaning. Therefore, information is not a mere synonymous of data.*

The ***Knowledge Economy*** *is defined as the production of products and services based on knowledge-intensive activities that contribute to an accelerated pace of technical and scientific advance, as well*

as rapid obsolescence. The key component of a knowledge economy is a greater reliance on intellectual capabilities than on physical inputs or natural resources.

*The **Internet Economy** is defined as covering the full range of our economic, social and cultural activities supported by the Internet and related information and communications technologies*[7]*.*

***Information or Knowledge workers** in the most basic definition are persons employed to produce or analyse ideas and information. Multiple sources define knowledge workers as workers creating knowledge capital, who process existing information to create new information to be used to define and solve problems. They include, as an example, medical practitioners, lawyers, judges, teachers, architects, engineers, managers or salespeople. Their main capital is knowledge, and they are mainly focused on "non-routine" tasks.*

***Data workers** collect, storage, manage and analyse data, as their primary activity. Data workers can be knowledge workers if they are focused on non-routine tasks. For example, data entry clerks' primary activity is related to data, so they are data workers. However, data entry is a very routine task and as such data entry clerks should not be considered as knowledge workers. Another category of data workers is data analysts, who usually extract and analyse information from one single source, such as a CRM database. They require a medium level of creative thinking and usually work on structured data.*

***Data scientists** require solid knowledge in statistical foundations and advanced data analysis methods combined with a thorough understanding of scalable data management, with the associated technical and implementation aspects (European Big Data Value Partnership Strategic Research and Innovation Agenda, April 2014). They can deliver novel algorithms and approaches such as advanced learning algorithms, predictive analytics mechanisms, etc. Data scientists should also have a deep knowledge of their businesses; the most difficult skills to find, include advanced analytics and predictive analysis skills, complex event processing skills, rule management skills, business intelligence tools, data integration skills (UNC, 2013).*

[7] "Measuring the Internet Economy: A Contribution to the Research Agenda", OECD Digital Economy Papers, 2013