# D2.1 FIRST REPORT ON FACTS AND FIGURES

# UPDATING THE EUROPEAN DATA MARKET MONITORING TOOL

**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**

**21st March 2018**

# Contents

# Table of Figures

## Table of Tables

## 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*[1]*.*

***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.*

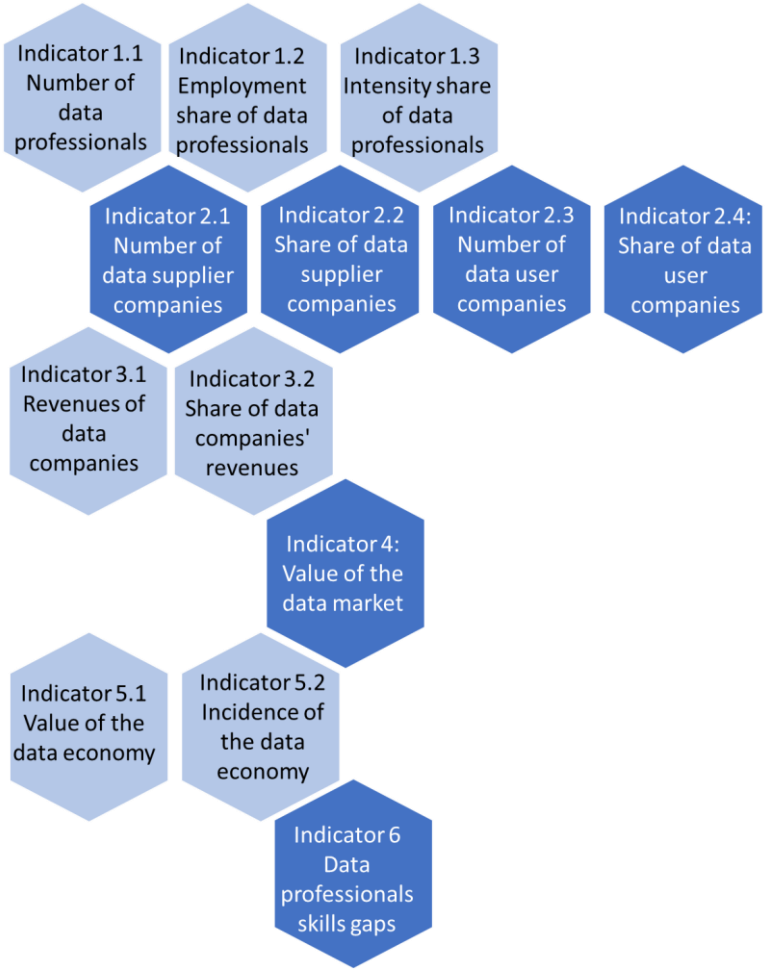
***Data workers** collect, storage, manage and analyse data, as their primary activity. Data workers can be knowledge workers if they are focused on non-routine tasks. For example, data entry clerks' primary activity is related to data, so they are data workers. However, data entry is a very routine 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).*

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

# EXECUTIVE SUMMARY

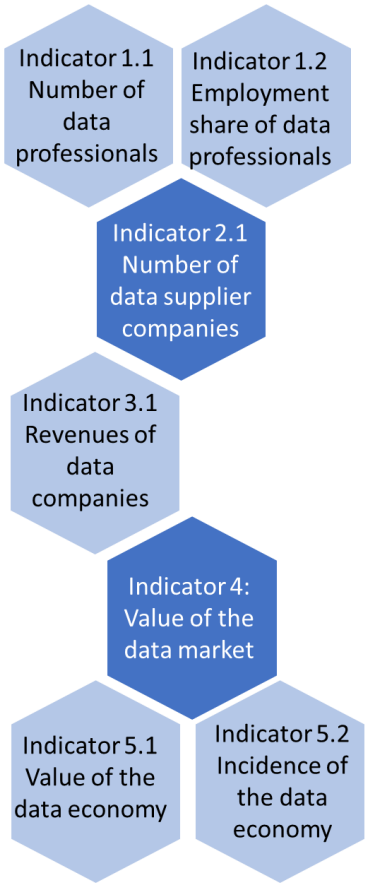
This First Report on Facts & Figures (Deliverable D2.1 of the Update of the European Data Market Study, SMART 2016/0063) presents the results obtained through the first round of measurements of the European Data Market Monitoring Tool, which was designed and developed under the previous European Data Market Study (SMART 2013/0063) and covered the period 2013-2015 with forecasts to the year 2020 under three distinct scenarios. In this update, the Monitoring Tool has been revised and extended to cover the years 2016-2017 with forecasts at 2025 under the same alternative scenarios. These 2025 scenarios take as a reference starting point the 2020 scenarios presented in February 2017 under the previous European Data Market Study (SMART 2013/0063). When presenting the forecasts of the indicators measured in this report, this document includes not only the results at 2025 but also those at 2020 under the Baseline scenario as previously measured. This is done to show the whole timeframe encompassed by the European Data Market Monitoring Tool as well as to ensure a high level of mutual intelligibility between the 2013 Study and the 2016 Study. This document focuses on the following set of indicators:




Each indicator is measured at the level of the total EU28 and for all 28 EU Member States, when available and applicable; industry-specific and company-size views are also offered with indicators provided by industry sector and company size bands, when possible.

Two different views of the data are presented at European level: the current EU28, and the EU27 excluding the United Kingdom.

As in the previous European Data Market Study (SMART 2013/0063), a selected number of indicators has been developed and updated for three non-European countries, namely Brazil, Japan and the United States. For each of these countries, this report presents the following indicators:




## Methodological Note

The measurement of each of the indicators is based on a sophisticated methodology that combines data collection, data modelling, and desk research.

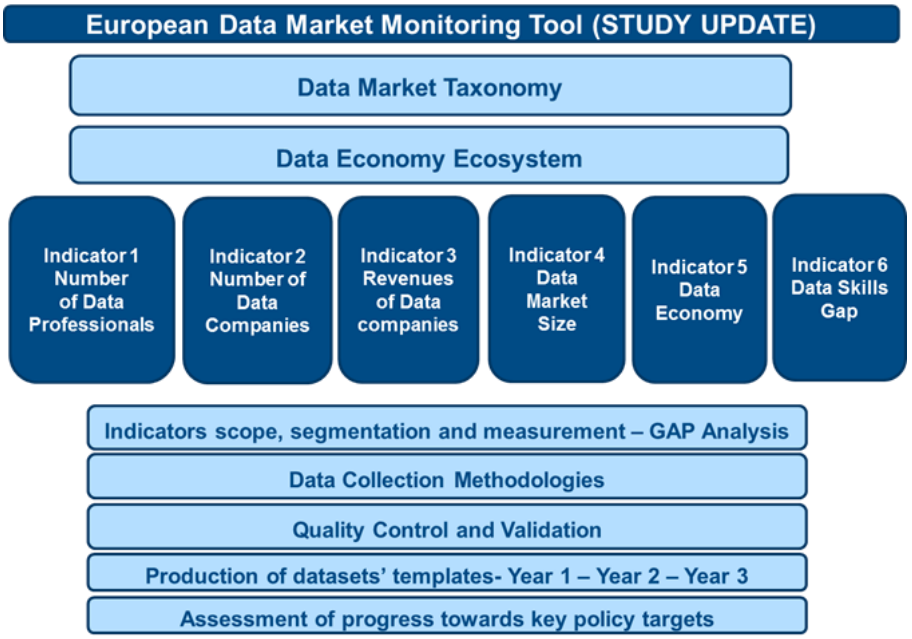
The survey conducted in 2015 among eight Member States and investigating the current and future role of data companies was confirmed as the basis for most of the key assumptions underpinning the status of the Data Market and the Data Economy in the period 2016-2017 and throughout 2025. The 2015 survey targeted potential data companies in two industries: ICT and Professional services, and data users in 11 industries. It has been complemented by a series of annual update surveys targeting all business sectors, and company-sizes greater than 10 employees. These surveys[2], originally conducted by IDC across a select number of Member States (usually the "Big Six" – France, Germany, Italy, Poland, Spain and the U.K.) have been regularly balanced to represent the mix of industries and size bands for companies across the whole of 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.

[2] Examples of additional surveys leverage in this report are: IDC European Vertical Market Survey conducted in September 2017; IDC Core IT Spending Guide, 2 Half 2016 conducted in July 2017; IDC Big Data and Analytics Spending Guide 2 Half 2016 conducted in August 2017. Additional details in Paragraph 1.4 below and in the Methodological Annex at the end of this report.

## Updating the European Data Market Monitoring Tool

Based on a modular and flexible structure, the European Data Market Monitoring Tool for the Update of the European Data Market Study leverages the existing tool that was used to measure the Data Market and the Data Economy during the period 2013-2016. 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.

*The updated European Data Market Monitoring Tool*




Source: IDC, 2017

## Three Developments Paths to 2025: The Baseline Scenario, the Challenge Scenario, the High-Growth Scenario.

As in the previous European Data Market Study (SMART 2013/0063), this update report presents three potential future scenarios of evolution of the European Data Market and Economy, driven by different macroeconomic and framework conditions. The scenarios presented here are focused on the year 2025, taking as a reference point the 2020 scenarios presented in February 2017[3]. The key assumptions of the 2020 scenarios have been revised and redefined for the 2025 scenarios. While the 2020 scenarios were mainly differentiated by different demand-supply dynamics, the 2025 scenarios are driven by the intersection of two main evolution paths (focal issues):

- 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. High power concentration means that a few leading data holders control most of the data assets; the opposite is an open and participatory social model of data sharing and management. In a centralised scenario, data

[3] "The European Data Market", Final report, February 2017, SMART 2013/0063, http://datalandscape.eu/study-reports

holders also tend to appropriate a high share of benefits from data innovation, while in a decentralised scenario the benefits are more equally shared.

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** (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.

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.

## Measuring the Data Professionals

***Data professionals***[4] *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 2016 and 2017*

All in all, data professionals are estimated at a total of 5.3 million in the EU27 and at 6.7 million in the EU28 in 2017, thus marking a significant increase in 2017 over the previous year (8.2% and 8% year-on-year respectively). When compared to the year 2017, the 2020 Baseline scenario outlined in the previous European Data Market Study (SMART 2013/0063) would register a Compound Annual Growth Rate (CAGR) of 7.9% and 7.5% at the level of EU27 and EU28 respectively. More interestingly, the employment share and the intensity share components of the data professionals' indicator are also expected to significantly improve in 2017 and 2020 if compared to our estimates in 2016 (now estimated at 3% and 3.4% in 2007 and 2020 in the EU27 and 3.2% and 3.6% for the same

[4] 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 update 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.

years in the EU28), thus confirming the positive evolution of the workforce involved in data-related professions over the period under consideration.

*Data Professionals, 2016-2017-2020 Baseline and Growth Rates*

| Indicator 1 — Data Professionals 2016-2017-2020 Baseline and Growth Rates | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| N. | Region | Name | Description | 2016 | 2017 | 2020 Baseline | Growth 2017/2016 | CAGR 2020/2017 |
| 1.1 | EU27 | Number of data professionals | Total number of data professionals in EU (000s) | **4,875** | **5,273** | **6,619** | **8.2%** | **7.9%** |
| 1.1 | EU28 | Number of data professionals | Total number of data professionals in EU (000s) | **6,187** | **6,685** | **8,309** | **8.0%** | **7.5%** |
| 1.2 | EU27 | Employment share of data professionals | Share of data professionals on total employment in EU (%) | **2.8%** | **3.0%** | **3.4%** | **5.3%** | **4.2%** |
| 1.2 | EU28 | Employment share of data professionals | Share of data professionals on total employment in EU (%) | **3.1%** | **3.2%** | **3.6%** | **5.3%** | **4.0%** |
| 1.3 | EU27 | Intensity share of data professionals | Average number of data professionals per user company (units) | **9.6** | **10.2** | **12.3** | **5.9%** | **6.3%** |
| 1.3 | EU28 | Intensity share of data professionals | Average number of data professionals per user company (units) | **9.2** | **9.7** | **11.5** | **5.8%** | **5.9%** |

## *Data Professionals at 2025*

The number of data professionals in both the EU27 and EU28 is forecast to grow significantly under all the new three scenarios at 2025 as the use of data-driven innovation is expected to grow under unabatedly even under the less economically favourable scenario.

A steady progression of the number of data professionals continues to emerge from our latest estimates. Under the new Baseline scenario, data professionals are expected to amount to 9.4 million in the EU27 and 11.5 million in the EU28 at 2025, thus representing a solid growth rate between 6.7% and 7.2% over the 2017-2025 period.

In the new Challenge and High Growth scenarios, data professionals would be more than 8.6 million in the EU27 and 10.5 million in the EU28 and 10.9 million and 13,4 million respectively. Under all new scenarios, the CAGR over the period 2017-2025 would be in line with the CAGR featured by the Data Market growth, thus confirming again the close relationship between the two variables.

*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) | 8,641 | 9,366 | 10,916 | 5.5% | 7.2% | 10.5% |
| 1.1 | EU28 | Number of data professionals | Total number of data professionals in EU (000s) | 10,487 | 11,477 | 13,450 | 4.8% | 6.7% | 10.1% |

## 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 2016 and 2017*

The number of data suppliers continues to grow at a faster pace than the number of data users: the former is estimated at almost 142,000 in the EU27 and at more than 276,000 units in the EU28, thus exhibiting a year-on-year growth of 5.7% in 2017. Data users, instead, are projected to grow at 2.1% in 2017, amounting to almost 516,000 in the EU27 and to 690,650 units in the EU28. If compared to the measurements carried out by the European Data Market Monitoring Tool over the period 2013-2016 these latest estimates show a more dynamic picture of data companies in the EU, with growth rates constantly increasing over the past four years.

This positive dynamic is reflected in the percentage shares 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 now estimated at 11.5% in the EU27 (up 0.6 percentage points vis-à-vis the

previous year) and 15% in the EU28 (up 0.8 percentage points), while the data users' penetration rates are up a modest 0.1 percentage point in 2017 in both the EU27 and EU28.

*Data Companies, 2016-2017-2020 Baseline and Growth Rates*

| Indicator 2 – Data Companies 2016-2017-2020 Baseline and Growth Rates | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| N. | Name | Description | Market | 2016 | 2017 | 2020 Baseline | Growth 2017/2016 | CAGR 2020/2017 |
| 2.1 | Number of data suppliers | Total number of data suppliers measured as legal entities based in the EU (000s) | EU27 | **134,300** | **141,900** | **157,150** | **5.7%** | **3.5%** |
| 2.1 | Number of data suppliers | Total number of data suppliers measured as legal entities based in the EU (000s) | EU28 | **261,450** | **276,450** | **305,600** | **5.7%** | **3.4%** |
| 2.2 | Share of data suppliers | % share of data companies on total companies in the ICT and Professional services industries | EU27 | **10.9%** | **11.5%** | **12.8%** | **5.6%** | **3.5%** |
| 2.2 | Share of data suppliers | % share of data companies on total companies in the ICT and Professional services industries | EU28 | **14.2%** | **15.0%** | **16.5%** | **5.5%** | **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** | **516,550** | **539,800** | **2.1%** | **1.5%** |
| 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** | **690,650** | **721,850** | **2.1%** | **1.5%** |
| 2.4 | Share of data users | % share of data users on total companies in the EU industry | EU27 | **5.7%** | **5.8%** | **6.1%** | **1.8%** | **1.6%** |
| 2.4 | Share of data users | % share of data users on total companies in the EU industry | EU28 | **6.5%** | **6.6%** | **6.9%** | **1.9%** | **1.6%** |

## *Data Suppliers Forecasts at 2025*

The outlook for data suppliers is continued growth, but the baseline growth to 2025 does not fully match the growth seen in 2020 according to our previous Baseline scenario – primarily because the market is larger and so growth becomes more difficult. In spite of 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 individuals 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.

*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 | 170,100 | 182,050 | 202,800 | 1.6% | 3.0% | 5.2% |
| EU28 | 338,700 | 357,950 | 410,900 | 2.1% | 3,2% | 6.1% |

### *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, as well as Education. 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 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 25 percent to 2025, compared with 1.7 % for those in the smaller size band, they do not add significantly to the total number of data companies.

*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 | 562,250 | 585,150 | 629,250 | 0.8% | 1.6% | 3.1% |
| EU28 | 758,100 | 786,650 | 853,200 | 1.0% | 1.7% | 3.4% |

## 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 2016 and 2017*

Data suppliers' revenues have increased by almost 11% in 2017 to reach 52.3 Billion Euros in the EU27 and 68.5 Billion Euro in the EU28 – a constant increase since 2013 according to our initial measurements and the Monitoring Tool. The share of the data companies' revenues on the total companies' revenues in the ICT and Professional services sectors has grown as a result and is now stable at 3.2% in both the EU27 and the EU28.

*Data Companies' Revenues and Growth, 2016-2017-2020 Baseline (€, Million; %)*

| Indicator 3 — Data Companies' Revenues and Growth | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| N. | Region | Name | Description | 2016 | 2017 | 2020 | Growth 2017/2016 | CAGR 2020/2017 |
| 3.1 | EU27 | Total revenues of data companies in the EU | Total revenues of the Data Suppliers calculated by Indicator 2 | 47,178 | 52,300 | 69,583 | 10.9% | 10.0% |
| 3.1 | EU28 | Total revenues of data companies in the EU | Total revenues of the Data Suppliers calculated by Indicator 2 | 61,781 | 68,575 | 89,280 | 11.0% | 9.2% |
| 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% | NA | 6.0% | 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% | 3.2% | NA | 3.4% | NA |

## *Data Companies' Revenues Forecasts at 2025*

Data 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 8.3% and of 7.3% from 2020 to 2025 - easily outpacing the growth of the total ICT market over the same period (expected to be 1.6% from 2016 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.

*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 | 77,077 | 96,921 | 133,372 | 2.1% | 6.9% | 13.9% |
| Total EU28 | 101,049 | 126,710 | 179,043 | 2.5% | 7.3% | 14.9% |

## 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 2016-2017*

The value of the Data Market in 2017 for both the EU27 and the EU28 is showing a buoyant growth rate of more than 9% year-on-year and is expected to surpass the threshold of 60 billion Euro in 2020 in the EU27 according to the Baseline scenario described in our previous study – 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 measures the Data Market at almost 46.2 billion Euro in 2016.

*Data Market Value and Growth, 2016-2017-2020 Baseline (€, Million; %)*

| Indicator 4 — Value and Growth of the Data Market | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| N. | Market | Name | Description | 2016 | 2017 | 2020 Baseline | Growth 2017/2016 | CAGR 2020 Baseline/2017 |
| 4.1 | EU27 | Value of the Data Market | Estimate of the overall value of the Data Market | **46,183** | **50,438** | **60,254** | **9.2%** | **6.1%** |
| 4.1 | EU28 | Value of the Data Market | Estimate of the overall value of the Data Market | **59,496** | **65,038** | **77,407** | **9.3%** | **6.0%** |

## *The Data Market Forecasts at 2025*

In 2025 the value of the Data Market under the new High Growth scenario is expected to largely double in size with respect to its 2017 estimates in both the EU27 and the EU28. This will correspond to a considerable CAGR for the period 2020-2025 of 12.7% and 13.6% in the EU27 and the EU28 respectively. According to our new 2025 Baseline scenario, the Data Market will amount to more than 85 billion Euro in the EU27, against 50.4 billion Euro in 2017 (a 7.1% CAGR 2020-2025), while under the Challenge scenario the Data Market will still represent 73.7 billion Euro, growing at a compound annual growth rate of 4.1% 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).

*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 | **73,677** | **85,078** | **109,633** | **4.1%** | **7.1%** | **12.7%** |
| Total EU28 | **96,098** | **110,057** | **146,433** | **4.4%** | **7.3%** | **13.6%** |

## 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 2016 and 2017*

The value of the Data Economy for the EU28 is to exceed the threshold of 300 billion Euro in 2016, and the estimated 2017 growth rate of 12% highlights a faster growth, considering the 7% for 2016. The share of overall impacts on GDP is expected to grow from 2.2% in 2016 to 2.4% in 2017. Results for EU27 and EU28 are similar, but it is worth highlighting the difference in the 2020/2017 CAGR and in 2017 growth rates for the two regions. EU27 rates are higher than EU28, highlighting uncertainty related to Brexit is affecting the U.K. which is showing lower than average growth rates and 2020/2017 CAGR.

*Data Economy Value and Growth, 2016-2017-2020 Baseline and Impacts on GDP 2016-2017 (€, Million; %)*

| Indicator 5 — Value and Growth of the Data Economy | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| N. | Market | Name | Description | 2016 | 2017 | 2020 Baseline | Growth 2017/2016 | CAGR 2020/2017 | Impacts on GDP 2016 | Impacts on GDP 2017 |
| 5.1<br>5.2 | EU27 | Value of the Data Economy and Impacts on EU GDP | Value of the Data Market's direct, indirect and induced impacts on the EU economy and % of EU GDP | **243,205** | **267,006** | **365,761** | **9.8%** | **11.1%** | **2.1%** | **2.2%** |
| 5.1<br>5.2 | EU28 | Value of the Data Economy and Impacts on EU GDP | Value of the Data Market's direct, indirect and induced impacts on the EU economy and % of EU GDP | **305,977** | **335,618** | **452,190** | **9.7%** | **10.4%** | **2.2%** | **2.4%** |

Source: European Data Market Monitoring Tool, IDC January 2018

### *The Data Economy Forecasts at 2025*

The Data Economy in 2025 is expected to near 700 billion Euro under the Baseline scenario impacting more than 4.2% over the EU28 and to arrive at 544 billion Euro in the EU27 (with a share on GDP of 5.4%)

The CAGR 2020/2025 in the EU28 is higher (17.2%) than the CAGR 2017/2025 (14.6%), thus highlighting an acceleration of growth after 2020, that will make the Data Economy for EU28 surpass the threshold of EUR 1 trillion, and accounting for 6% of the GDP at 2025. Similarly, the Challenge scenario will see a slowdown of the economic effects, from 4.3% CAGR 2017/2025 to 0.8% CAGR 2020/2025, with the Data Economy being well below 500 billion Euro.

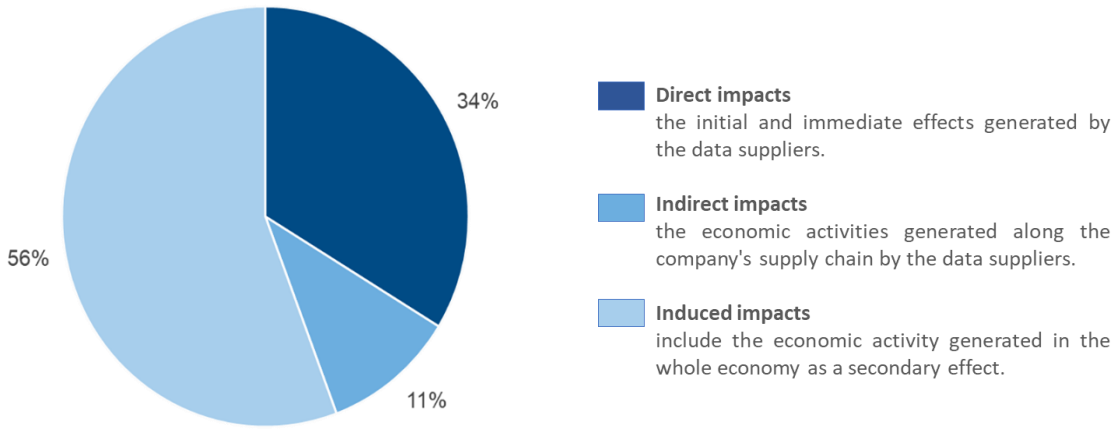
*Data Economy Forecast in 2025 and Impacts on GDP according to the Three Scenarios (€, Million; %)*

| 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<br>5.2 | Value of the Data Economy<br>EU27 | Value of the Data Market plus estimate of direct, indirect and induced impacts on the economy | **376,520** | **544,241** | **769,505** | **2.9%** | **4.0%** | **5.4%** |
| 5.1<br>5.2 | Value of the Data Economy<br>EU28 | Value of the Data Market plus estimate of direct, indirect and induced impacts on the economy | **470,402** | **669,197** | **1,001,073** | **3.0%** | **4.2%** | **6.0%** |

Source: European Data Market Monitoring Tool, IDC January 2018

As in the previous study, this report provides a detailed insight of the Data Economy by type of impact – direct, indirect and induced impacts. The composition of impacts changes along time, from 2017 (figure below) to 2025, in favour of induced impacts, this 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 56% (both in EU27 and in EU28) from around 30% in 2017. Indirect impacts in turn will lose around 11% of share, but still in 2025 accounting for a very high percentage (34%). With respect to 2017, 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.

*Data Economy by Type of Impact, EU28, 2025 (%)*




Source: European Data Market Monitoring Tool, IDC January 2018

## 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. In the year 2017, given the strong increase of demand from 2016, the estimated gap reached approximately 449,000 unfilled positions, corresponding to 6% of total demand. By 2020 we expect the gap to expand to 1 million unfilled positions in the EU27.

The data skills gap will continue to grow in all 3 forecast scenarios, reaching 1 million unfilled positions in 2025 in the EU27 Baseline scenario (11% of demand) but up to over 2 million in the EU27 High Growth scenario (19% of demand). In the Challenge scenario the 2025 gap is slightly lower than in the Baseline. The main trends explaining these variations are the following:

- Gradual increase of the number of graduates in the period 2020-2025, thanks to increasing awareness of the market potential and to the policies promoting STEM and data skills education. These policies were launched in several countries since 2015 and their effects are starting to be seen in the forecast period. For example, Ireland has already seen an increase of 54% of level 8 graduates in the STEM field from 2012 to 2018[5];
- Increase of inflows in the data skills market from other careers, upskilling and re-training initiatives, due to higher attractiveness of the ICT career in the Baseline and even more so in the High Growth scenario;
- In the Challenge scenario, lower inflows to the data skills market from other careers, upskilling and re-training initiatives, because of the slower development of the Data Market;

5 Presentation at the workshop "Forecasting ICT high-level skills demand", on the ICT skills Action Plan 2014-2018, by Higher Education Policy and Research Department of Education and Skills, Irish Government, 2018

- In all 3 scenarios, fast growth of data professionals demand that overcomes supply growth.

*Data Professionals Skills Gap in the EU, 2016-2017-2020 Baseline and 2025 - Three Scenarios*

| Indicator 6 — Data Professionals Skills Gap | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| N. | Name | Description | | Actual | | Baseline | Baseline Scenario | | Challenge Scenario | | High Growth Scenario | |
| | | | | 2016 | 2017 | 2020 | 2025 | 17-25 CAGR | 2025 | 17-25 CAGR | 2025 | 17-25 CAGR |
| 6.1 | Data professionals skills gap | Gap between demand and supply of data professionals, N, 000s | EU27 | 343 | 378 | 603 | 1,007 | 13% | 904 | 11.5% | 2,116 | 24% |
| | | | EU28 | 428 | 449 | 699 | 1,111 | 12% | 992 | 10.4% | 2,393 | 23.3% |
| 6.2 | | Gap between demand and supply of data professionals, % | EU27 | 6.2% | 6% | 9% | 11% | | 10.4% | | 19% | |
| | | | EU28 | 6.2% | 6% | 8% | 10% | | 9.4% | | 18% | |

Source: European Data Market Monitoring Tool, IDC 2018

## 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, marks the strongest growth in 2017 vis-à-vis the other two international partners under consideration (10% and 8.7% year-on-year growth in the number of data professionals and employment share in 2017 over 2016 respectively). The same applies for the data supplier companies' indicators, with the highest increase of data suppliers in 2017 (4.6% vs. 1.1% and 3% in Brazil and Japan respectively) and an even stronger rise in their associated revenues and in the value of the Data Market in 2017 over 2016 (12.7% vs. 4.3% and 8.7% in Brazil and Japan respectively). Accordingly, the U.S.' overall Data Economy (direct and backward indirect impacts only) has increased by almost 7% in 2017 gaining a 4.3% year-on-year in incidence on GDP and now representing more than 0.8% of the country's GDP.

*USA Indicators - Overview 2016, 2017*

| USA – Indicators' Overview | | | | | |
|---|---|---|---|---|---|
| N. | Name | Metrics | 2016 | 2017 | Growth 2017/2016 |
| 1.1 | Number of Data professionals | Total Number of Data professionals (Thousands) | 12,732 | 14,012 | 10.05% |
| 1.2 | Data professionals' | % of Data professionals on | 8.42% | 9.15% | 8.66% |

| USA – Indicators' Overview | | | | | |
|---|---|---|---|---|---|
| N. | Name | Metrics | 2016 | 2017 | Growth 2017/2016 |
| | employment share | total employment | | | |
| 2.1 | Number of Data Suppliers | Total number of data supplier companies (000s) | 289,556 | 302,810 | 4.58% |
| 3.1 | Revenues of Data Companies | Total revenues generated by companies specialized in the supply of data-related products and services (Million €) | 129,173 | 145,546 | 12.67% |
| 4.1 | Value of the Data Market | Estimate of the overall a value of the Data Market (Million €) | 129,173 | 145,546 | 12.67% |
| 4.2 | Value of the Data Economy (Only Direct and Backward Indirect impacts) | Direct Impacts (Million €) | 108,521 | 113,677 | 4.75% |
| | | Backward Indirect Impacts (Million €) | 7,270 | 7,766 | 6.82% |
| 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% | 0.81% | 4.25% |

Source: European Data Market Monitoring Tool, IDC 2018

### *Brazil*

Brazil has three years of weak growth, with 2015 and 2016 showing GDP declines, according to the IMF GDP data. 2017 show more promise, with GDP growth forecast, but there is momentum to overcome from the three low growth years, and investment is weak. However, the growth in data professionals, data companies, and the Data Economy is inevitable because of the benefits that come from digitally transforming companies. Brazil outlook is more positive for 2017, but the country still has some catching up to do when compared with European investment in the Data Economy.

Data professionals and their associated share on total employment in Brazil have grown between 1.4% and 2.1%, the number of data supplier companies, their related revenues and the accompanying Data Market have grown of a considerable 4.3%. As a result, the incidence of the economy on Brazil's GDP will also increase but only of a modest 0.4% year-on-year (with the Data Economy – direct and backward indirect impacts – representing just 0.16% of Brazil's GDP in 2017). See Figure below for an overview of Brazil's growth rates in 2017.

*Brazil Indicators - Overview 2016, 2017*

| Brazil – Indicators' Overview | | | | | |
|---|---|---|---|---|---|
| N. | Name | Metrics | 2016 | 2017 | Growth rate 2017/2016 |
| 1.1 | Number of Data professionals | Total Number of Data professionals (Thousands) | 1,160 | 1,176 | 1.37% |
| 1.2 | Data professionals' employment share | % of Data professionals on total employment | 1.81% | 1.84% | 2.13% |
| 2.1 | Number of Data Suppliers | Total number of data supplier companies (000s) | 35,979 | 36,387 | 1.13% |
| 3.1 | Revenues of Data Companies | Total revenues generated by companies specialized in the supply of data-related products and services (Million €) | 6,049 | 6,310 | 4.31% |
| 4.1 | Value of the Data Market | Estimate of the overall a value of the Data Market (Million €) | 6,049 | 6,310 | 4.31% |
| 4.2 | Value of the Data Economy (Only Direct and Backward Indirect impacts) | Direct Impacts (Million €) | 6,157 | 6,395 | 3.86% |
| | | Backward Indirect Impacts (Million €) | 290 | 298 | 2.72% |
| 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% | 0.16% | 0.38% |

Source: European Data Market Monitoring Tool, IDC 2018

## *Japan*

The indicators measuring the state of the Data Market and the Data Economy in Japan have all registered a significant growth in 2017 over the previous year. The number and employment share of data professionals rose by 8% and 4% respectively in 2017, while the number of data suppliers increased by 3% generating a significant growth in revenues and Data Market of 8.6%. The Data Economy has therefore also marked a positive development in Japan with an incidence on GDP now at 0.95% of the country's GDP – a growth of 1.6% with respect to the incidence in 2016. The Figure below presents the main growth rates for Japan's Data Economy.

*Japan Indicators - Overview 2016, 2017*

| Japan – Indicators' Overview | | | | | |
|---|---|---|---|---|---|
| N. | Name | Metrics | 2016 | 2017 | Growth rate 2017/2016 |
| 1.1 | Number of Data professionals | Total Number of Data professionals (Thousands) | 3,740 | 4,040 | 8.03% |
| 1.2 | Data professionals' employment share | % of Data professionals on total employment | 5.82% | 6.05% | 4.00% |
| 2.1 | Number of Data Suppliers | Total number of data supplier companies (000s) | 101,612 | 104,664 | 3.00% |
| 3.1 | Revenues of Data Companies | Total revenues generated by companies specialized in the supply of data-related products and services (Million €) | 25,513 | 27,723 | 8.66% |
| 4.1 | Value of the Data Market | Estimate of the overall a value of the Data Market (Million €) | 25,513 | 27,723 | 8.66% |
| 4.2 | Value of the Data Economy (Only Direct and Backward Indirect impacts) | Direct Impacts (Million €) | 27,394 | 29,949 | 9.33% |
| | | Backward Indirect Impacts (Million €) | 1,189 | 1,269 | 6.66% |

| Japan – Indicators' Overview | | | | | |
|---|---|---|---|---|---|
| N. | Name | Metrics | 2016 | 2017 | Growth rate 2017/2016 |
| 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% | 0.95% | 1.59% |

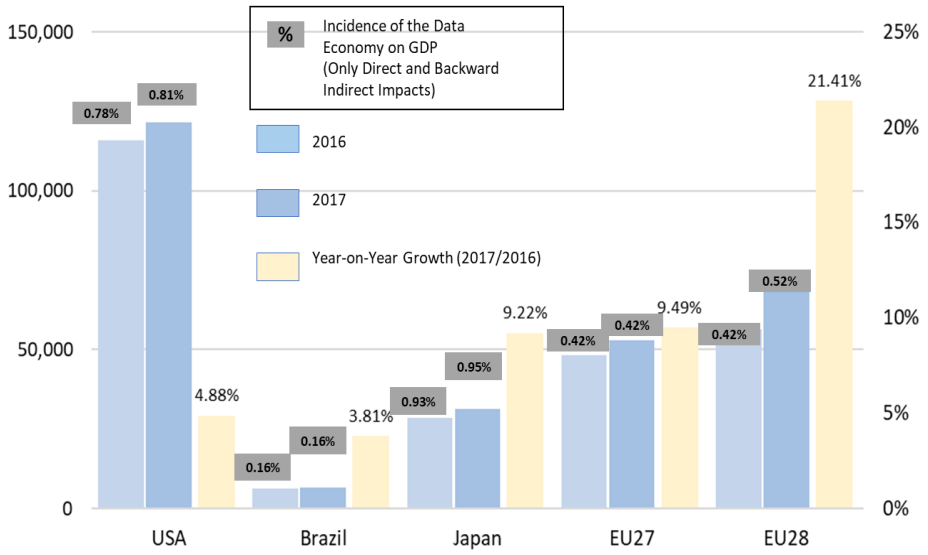
Source: European Data Market Monitoring Tool, IDC 2018

### *International Overview and Comparison with the EU*

In line with the results obtained in the previous European Data Market Study (SMART 2013/0063), the U.S. continue to enjoy the highest impact of the Data Economy on their GDP - 0.81% in 2017, up 4.9% with respect to 2016.

While not leading in absolute values, 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 represents a growing and dynamic 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.

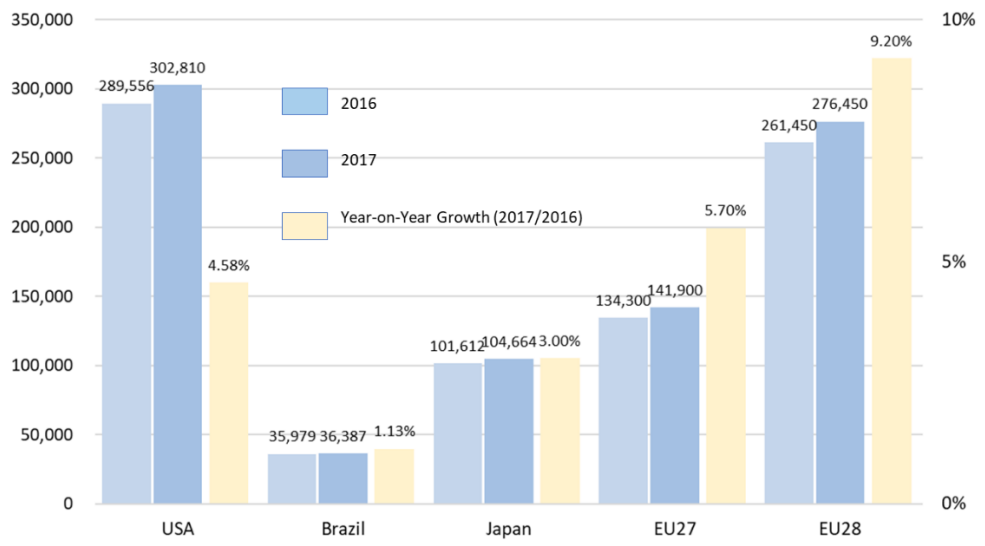
*Value, Growth and Incidence of the Data Economy as a Percentage of GDP in the U.S., Brazil, Japan and EU, 2016- 2017 (€, Million, %)*




Source: European Data Market Monitoring Tool, IDC 2018

Interestingly, though, Europe appears to be catching up on its gap with the U.S. and shows a renovated dynamism in some of the most significant Data Economy's areas. In terms of data suppliers, for example, the EU can exhibit a year-on-year growth 2017-2016 of 9.2% - more than twice than in the U.S. and three times stronger than in Japan over the same period.

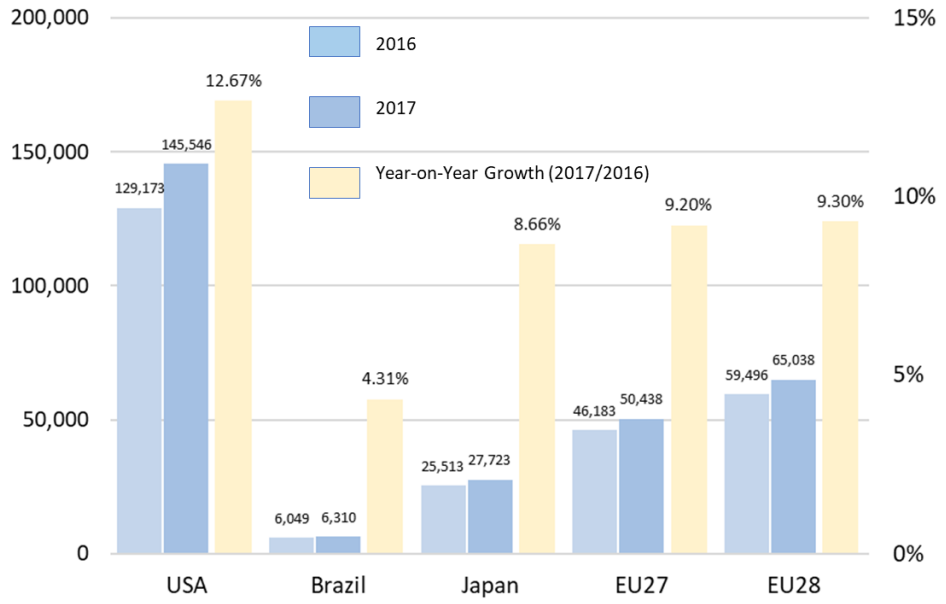
*Number of Data Suppliers in the U.S., Brazil, Japan and EU, 2016-2017, Growth 2017 (Units; '000, %)*




Source: European Data Market Monitoring Tool, IDC 2018

While close to double-digit growth, the value of the Data Market in the EU28 has marked a relative halt with respect to the U.S. in the period 2016-2017 but has continued to largely outpace Brazil and, to a much lesser extent, Japan.

*Value and Growth of the Data Market in the U.S., Brazil, Japan and EU, 2016-2017, Growth 2017 (Units; '000, %)*




Source: European Data Market Monitoring Tool, IDC 2018

Notwithstanding its dynamism, the EU continues to suffer from higher levels of fragmentation: Europe is more divergent than the U.S. when it comes to the usage of digital technologies across companies, including those in the same sector. This is partly due to structural factors (such as the higher presence of SMEs in Europe than in the U.S.) to cultural and educational factors (such as Europe's relatively weaker position in creating and keeping the necessary digital skills to support the digital transformation process).

Nevertheless, the EU remains a protagonist in many areas of the Data Economy - it has a lively digital startups scene and a considerable innovation capacity but it is still unable to translate this potential into global digital platforms as it is done in the U.S. - bar few notable exceptions, to be fair, such as Spotify in Sweden, Deezer in France, Shazam in the U.K. and Gemalto in the Netherlands, just to mention a few. In terms of research & innovation, and of the subsequent innovative technologies, Europe is slower the U.S. in building an effective ecosystem to turn these technologies in commercially exploitable applications. In areas such as Robotics, Augmented Reality/Virtual Reality and machine learning for instance – all future engines of the digital transformation – companies like ABB, BMW, Bosch and Siemens are investing heavily but find it difficult to build a critical mass competing at a par with American manufacturers.

# 1. INTRODUCTION

In 2013 the European Commission contracted IDC and Open Evidence to conduct a comprehensive study on the size and trends of the EU Data Market and Data Economy. The ensuing European Data Market Study (SMART 2013/0063) designed, developed and implemented a European Data Market Monitoring Tool with the objective to provide facts and figures on the size and trends of the EU Data Market and Data Economy.

The tool measured a series of indicators apprehending Europe's Data Market and Data Economy along several axes: from the type of workforce involved in the Data Market and Data Economy, to the companies supplying and using data for the production, distribution and sale of data-related product and services; from the potential gap emerging between the demand and supply of data workers, to the level of citizens' reliance in data-related product and services; from the size of the market where data-related products and services are exchanged, to the impacts that such market is exerting on the whole of Europe's economy (the Data Economy).

The tool developed and measured the above-mentioned set of indicators for the years 2013 through 2016 as well as presented three potential scenarios of evolution of the European Data Market and Data Economy to 2020, based on alternative development paths that were modelled on possible different evolutions of a set of key underpinning factors (such as the macroeconomic factors, policy and regulatory conditions, supply-demand dynamics of the market of data and global megatrends).

## 1.1 Objectives

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 has commissioned an update of the European Data Market (EDM) Study. 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 Overview of the Indicators

This document focuses on the following set of indicators:

- Indicator 1.1 Number of data professionals;
- Indicator 1.2 Employment share of data professionals;
- Indicator 1.3 Intensity share of data professionals;
- Indicator 2.1: Number of data supplier companies;
- Indicator 2.2: Share of data supplier companies;
- Indicator 2.3: Number of data user companies;
- Indicator 2.4: Share of data user companies;
- Indicator 3.1: Revenues of data companies;
- Indicator 3.2: Share of data companies' revenues;
- Indicator 4: Value of the Data Market;
- Indicator 5.1: Value of the Data Economy;
- Indicator 5.2: Incidence of the Data Economy;
- Indicator 6: Data professionals' skills gaps.

Each indicator is measured at the level of the total EU28 and for all 28 EU Member States, when available and applicable; industry-specific and company-size views are also offered with indicators provided by industry sector and company size bands, when possible.

Two different views of the data are presented at European level: the current EU28, and the EU27 excluding the United Kingdom.

As in the first 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. For each of these countries, this report presents the following indicators:

- Indicator 1.1: Number of data professionals;
- Indicator 1.2: Data professionals' employment share;
- Indicator 2.1: Number of data supplier companies;
- Indicator 3.1: Revenues of data companies;
- Indicator 4: Value of the Data Market;
- Indicator 5.1: Value of the Data Economy (only Direct and Backward Indirect impacts);
- Indicator 5.2: Incidence of the Data Economy on GDP (only direct and backward indirect impacts).

The table below offers an overview of the full set of indicators that have been developed in this First Report on Facts and Figures.

*Table 1: D2.1: First Report on Facts and Figures – Full Set of Indicators provided*

| # | Name of Indicator | Year | Industry | Member State | Company Size | EU28 Level | EU27 Level |
|---|---|---|---|---|---|---|---|
| 1.1 | Number of data professionals | 2016 | Delivered | Delivered | Not applicable | Delivered | Delivered |
| | | 2017 | Delivered | Delivered | Not applicable | Delivered | Delivered |
| | | 2020 Baseline | Delivered | Delivered | Not applicable | Delivered | Delivered |
| | | 2025 Baseline | Not applicable | Delivered | Not applicable | Delivered | Delivered |
| | | 2025 Challenge | Not applicable | Delivered | Not applicable | Delivered | Delivered |
| | | 2025 High Growth | Not applicable | Delivered | Not applicable | Delivered | Delivered |
| 1.2 | Employment Share | 2016 | Delivered | Delivered | Not applicable | Delivered | Delivered |
| | | 2017 | Delivered | Delivered | Not applicable | Delivered | Delivered |
| | | 2020 Baseline | Delivered | Delivered | Not applicable | Delivered | Delivered |
| | | 2025 Baseline | Not applicable | Not applicable | Not applicable | Not applicable | Not applicable |
| | | 2025 Challenge | Not applicable | Not applicable | Not applicable | Not applicable | Not applicable |
| | | 2025 High Growth | Not applicable | Not applicable | Not applicable | Not applicable | Not applicable |
| 1.3 | Intensity Share | 2016 | Delivered | Delivered | Not applicable | Delivered | Delivered |
| | | 2017 | Delivered | Delivered | Not applicable | Delivered | Delivered |
| | | 2020 Baseline | Delivered | Delivered | Not applicable | Delivered | Delivered |
| | | 2025 Baseline | Not applicable | Not applicable | Not applicable | Not applicable | Not applicable |
| | | 2025 Challenge | Not applicable | Not applicable | Not applicable | Not applicable | Not applicable |
| | | 2025 High Growth | Not applicable | Not applicable | Not applicable | Not applicable | Not applicable |
| 2.1 | Number of Data Suppliers | 2013 | Delivered | Delivered | Delivered | Delivered | Delivered |
| | | 2016 | Delivered | Delivered | Delivered | Delivered | Delivered |
| | | 2017 | Delivered | Delivered | Delivered | Delivered | Delivered |
| | | 2020 Baseline | Delivered | Delivered | Delivered | Delivered | Delivered |
| | | 2025 Baseline | Delivered | Delivered | Delivered | Delivered | Delivered |

| # | Name of Indicator | Year | Industry | Member State | Company Size | EU28 Level | EU27 Level |
|---|---|---|---|---|---|---|---|
| | | 2025 Challenge | Delivered | Delivered | Delivered | Delivered | Delivered |
| | | 2025 High Growth | Delivered | Delivered | Delivered | Delivered | Delivered |
| 2.2 | Share of Data Suppliers | 2016 | Delivered | Delivered | Delivered | Delivered | Delivered |
| | | 2017 | Delivered | Delivered | Delivered | Delivered | Delivered |
| | | 2020 Baseline | Delivered | Delivered | Delivered | Delivered | Delivered |
| | | 2025 Baseline | Not applicable | Not applicable | Not applicable | Not applicable | Not applicable |
| | | 2025 Challenge | Not applicable | Not applicable | Not applicable | Not applicable | Not applicable |
| | | 2025 High Growth | Not applicable | Not applicable | Not applicable | Not applicable | Not applicable |
| 2.3 | Number of Data Users | 2016 | Delivered | Delivered | Delivered | Delivered | Delivered |
| | | 2017 | Delivered | Delivered | Delivered | Delivered | Delivered |
| | | 2020 Baseline | Delivered | Delivered | Delivered | Delivered | Delivered |
| | | 2020 Baseline | Delivered | Delivered | Delivered | Delivered | Delivered |
| | | 2025 Baseline | Delivered | Delivered | Delivered | Delivered | Delivered |
| | | 2025 Challenge | Delivered | Delivered | Delivered | Delivered | Delivered |
| | | 2025 High Growth | Delivered | Delivered | Delivered | Delivered | Delivered |
| 2.4 | Share of Data Users | 2016 | Delivered | Delivered | Delivered | Delivered | Delivered |
| | | 2017 | Delivered | Delivered | Delivered | Delivered | Delivered |
| | | 2020 Baseline | Delivered | Delivered | Delivered | Delivered | Delivered |
| | | 2025 Baseline | Delivered | Delivered | Delivered | Delivered | Delivered |
| | | 2025 Challenge | Delivered | Delivered | Delivered | Delivered | Delivered |
| | | 2025 High Growth | Delivered | Delivered | Delivered | Delivered | Delivered |
| 3.1 | Revenues of Data Companies | 2016 | Not applicable | Delivered | Delivered | Delivered | Delivered |
| | | 2017 | Not | Delivered | Delivered | Delivered | Delivered |

| # | Name of Indicator | Year | Industry | Member State | Company Size | EU28 Level | EU27 Level |
|---|---|---|---|---|---|---|---|
| | | | applicable | | | | |
| | | 2020 Baseline | Not applicable | Delivered | Delivered | Delivered | Delivered |
| | | 2025 Baseline | Not applicable | Delivered | Delivered | Delivered | Delivered |
| | | 2025 Challenge | Not applicable | Delivered | Delivered | Delivered | Delivered |
| | | 2025 High Growth | Not applicable | Delivered | Delivered | Delivered | Delivered |
| **3.2** | Share of Data Companies' Revenues | 2016 | Delivered | Delivered | Delivered | Delivered | Delivered |
| | | 2017 | Delivered | Delivered | Delivered | Delivered | Delivered |
| | | 2020 Baseline | Not applicable | Not applicable | Not applicable | Not applicable | Not applicable |
| | | 2025 Baseline | Not applicable | Not applicable | Not applicable | Not applicable | Not applicable |
| | | 2025 Challenge | Not applicable | Not applicable | Not applicable | Not applicable | Not applicable |
| | | 2025 High Growth | Not applicable | Not applicable | Not applicable | Not applicable | Not applicable |
| **4** | Value of the Data Market | 2016 | Delivered | Delivered | Not applicable | Delivered | Delivered |
| | | 2017 | Delivered | Delivered | Not applicable | Delivered | Delivered |
| | | 2020 Baseline | Delivered | Delivered | Not applicable | Delivered | Delivered |
| | | 2025 Baseline | Delivered | Delivered | Not applicable | Delivered | Delivered |
| | | 2025 Challenge | Delivered | Delivered | Not applicable | Delivered | Delivered |
| | | 2025 High Growth | Delivered | Delivered | Not applicable | Delivered | Delivered |
| **5.1** | Value of the Data Economy | 2016 | Not applicable | Delivered | Not applicable | Delivered | Delivered |
| | | 2017 | Not applicable | Delivered | Not applicable | Delivered | Delivered |
| | | 2020 Baseline | Not applicable | Delivered | Not applicable | Delivered | Delivered |
| | | 2025 Baseline | Not applicable | Delivered | Not applicable | Delivered | Delivered |
| | | 2025 Challenge | Not applicable | Delivered | Not applicable | Delivered | Delivered |

| # | Name of Indicator | Year | Industry | Member State | Company Size | EU28 Level | EU27 Level |
|---|---|---|---|---|---|---|---|
| | | 2025 High Growth | Not applicable | Delivered | Not applicable | Delivered | Delivered |
| 5.2 | Incidence of the Data Economy on GDP | 2016 | Not applicable | Delivered | Not applicable | Delivered | Delivered |
| | | 2017 | Not applicable | Delivered | Not applicable | Delivered | Delivered |
| | | 2020 Baseline | Not applicable | Delivered | Not applicable | Delivered | Delivered |
| | | 2025 Baseline | Not applicable | Delivered | Not applicable | Delivered | Delivered |
| | | 2025 Challenge | Not applicable | Delivered | Not applicable | Delivered | Delivered |
| | | 2025 High Growth | Not applicable | Delivered | Not applicable | Delivered | Delivered |
| 6 | Data professional skills gap | 2016 | Not applicable | Delivered | Not applicable | Delivered | Delivered |
| | | 2017 | Not applicable | Delivered | Not applicable | Delivered | Delivered |
| | | 2020 Baseline | Not applicable | Delivered (Big Six) | Not applicable | Delivered | Delivered |
| | | 2025 Baseline | Not applicable | Delivered (Big Six) | Not applicable | Delivered | Delivered |
| | | 2025 Challenge | Not applicable | Delivered (Big Six) | Not applicable | Delivered | Delivered |
| | | 2025 High Growth | Not applicable | Delivered (Big Six) | Not applicable | Delivered | Delivered |

This document is accompanied by a comprehensive dataset reporting the values for all the indicators outlined in the table above, including the values obtained through the previous rounds of measurement of the Data Market and the Data Economy that were presented in the European Data Market Study (SMART 2013/0063).

## 1.3 Structure of the Document

This First Report on Facts & Figures is focused on the results obtained through the first round of measurements of the European Data Market Monitoring Tool for the period 2016-2017 with forecasts at 2025 under three distinct scenarios. The report is organized along the following chapters:

- Chapters 1 and 2 include a brief introduction and a short reminder of the overall study's goals and objectives as well as a summary of the European Data Market Monitoring Tool and its functioning.
- Chapter 3 covers the three evolution scenarios of the Data Market and the methodological approach to the indicators' forecasts, with specific reference to the assumptions' changes

that have been introduced to produce this Update of the European Data Market Study and the extension of the forecast period to the year 2025.

- Chapter 4 is devoted to the measurement of the data professionals including their main values in absolute terms, their share in terms of total employment and their forecast to 2025 according to the three scenarios under consideration.
- Chapters 5 and 6 provide the values for the indicators measuring the data companies (both suppliers and users of data) in terms of absolute numbers and produced revenues. It includes the updated forecast of the indicator by 2025.
- Chapter 7 presents the indicators measuring the size of the Data Market in Europe based on the total spending on software, hardware, and IT services' technologies and its contribution to the Data Market. It projects the forecast of the indicator by 2025.
- Chapter 8 measures the Data Economy in Europe for the years 2016 and 2017, as well as the forecasts to 2025 according to the three scenarios under consideration.
- Chapter 9 is devoted to the update of the data professionals' skills gap indicator in the EU.
- Chapter 10, the Data Economy beyond the EU, presents a select number of indicators for Brazil, Japan and the United States.
- Chapter 11 provides a set of concluding remarks of the report.
- The Methodological Annex summarises the key methodological steps that we have undertaken to measure the indicators covered in both the previous reports and in the current report.

## 1.4 Methodology Approach

The measurement of each of the indicators is based on a sophisticated methodology that combines data collection, models, and desk research. Some initial assumptions are built on surveys completed during March 2015, which are supported by ongoing annual surveys. The 2015 survey includes 8 Member States, and the annual surveys include 6 Member States.

The initial survey targeted potential data companies in two industries (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 survey is 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.

Data sources, their use, and date updated are as follows:

| Data Source | Updated | Used in |
|---|---|---|
| Eurostat Business Demographic Statistics | Dec 2017 | Data professionals<br>Data companies<br>Data users |
| Eurostat annual structural business statistics | Dec 2017 | Data professionals<br>Data companies<br>Data users |
| Eurostat chain linked Volumes (GDP) | Dec 2017 | Data Market<br>Data Revenues |
| IDC Core IT Spending guide 2H2016 | Jul 2017 | Data Market<br>Data Revenues |
| IDC Worldwide Black Book (standard edition) | Nov 2017 | Data Market<br>Data professionals |

| Data Source | Updated | Used in |
|---|---|---|
| | | Data Companies<br>Data Users<br>Data Revenues |
| IDC European Vertical Markets survey (2017) | Sep 2017 | Data Market |
| IMF World Economic Outlook | Oct 2017 | Data Market<br>Data Revenues<br>Data Economy |
| Consensus Forecasts – Consensus economics | Nov 2017 | Data Market<br>Data Revenues<br>Data Economy |
| IT Big Data and Analytics spending Guide 2H2016 | Aug 2017 | Data Market |
| ILOSTAT statistics and databases | Jan 2018 | Data Professionals |

Additional relevant sources leveraged in this report were IDC annual Small and Medium Business and Vertical Markets end user surveys and IDC End-User IT Trends and Digital Transformation: IDC European Vertical Markets Survey 2017 whose results were used to confirm and adjust estimates, when necessary, of the number of companies that were data users and data suppliers. The detailed data companies survey from 2017 provided a solid baseline for this estimate, and the annual end-user survey by size and vertical market identified any notable changes from 2015 and 2016. IDC's end-user survey asks specific questions about the actual and planned adoption of Big Data and Analytics, which gives a clear indication of trends in data use and supply.

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 and Citizens' Reliance on the Data Market for the year 2016, 2017 and for the three 2025 scenarios.

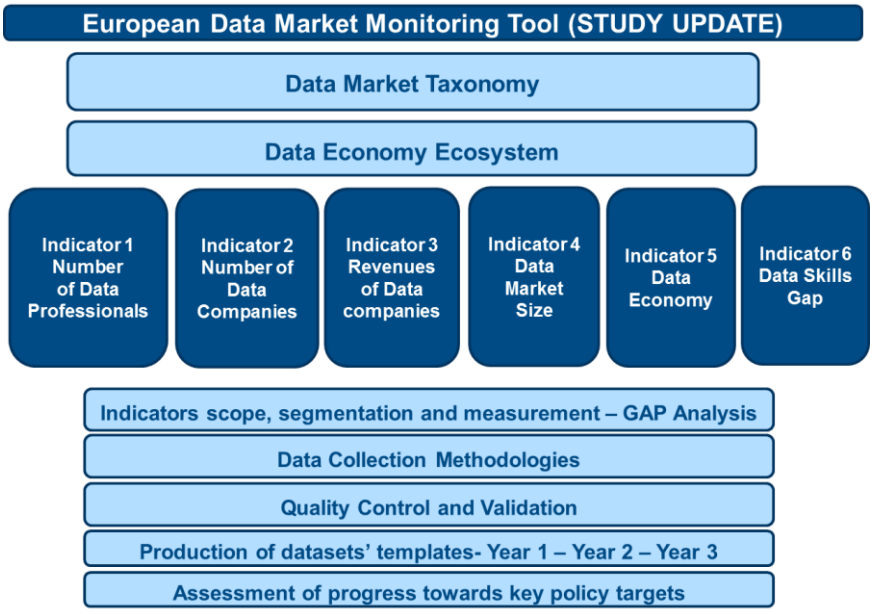
A comprehensive and detailed description of this report's methodology approach is offered in the Methodology Annex at the end of the document.

# 2. UPDATING THE EUROPEAN DATA MARKET MONITORING TOOL

## 2.1 The European Data Market Monitoring Tool

Based on a modular and flexible structure, the European Data Market Monitoring Tool for the Update of the European Data Market Study leverages the existing tool that was used to measure the Data Market and the Data Economy during the period 2013-2016. 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 1: The Updated EDM Monitoring Tool*




Source: IDC, 2017

## 2.2 Data Economy Ecosystem

The top layers of the EDM Tool describe the conceptual framework of the Data Market with the main definitions including:

- **Data Market Taxonomy** providing clear definition of all the main terms used in the analysis and in the Monitoring Tool, providing an objective and scientific basis for the definition and scope of the indicators.
- **Data Economy Ecosystem,** a conceptual framework describing the main components of the Data Market and their interrelations.

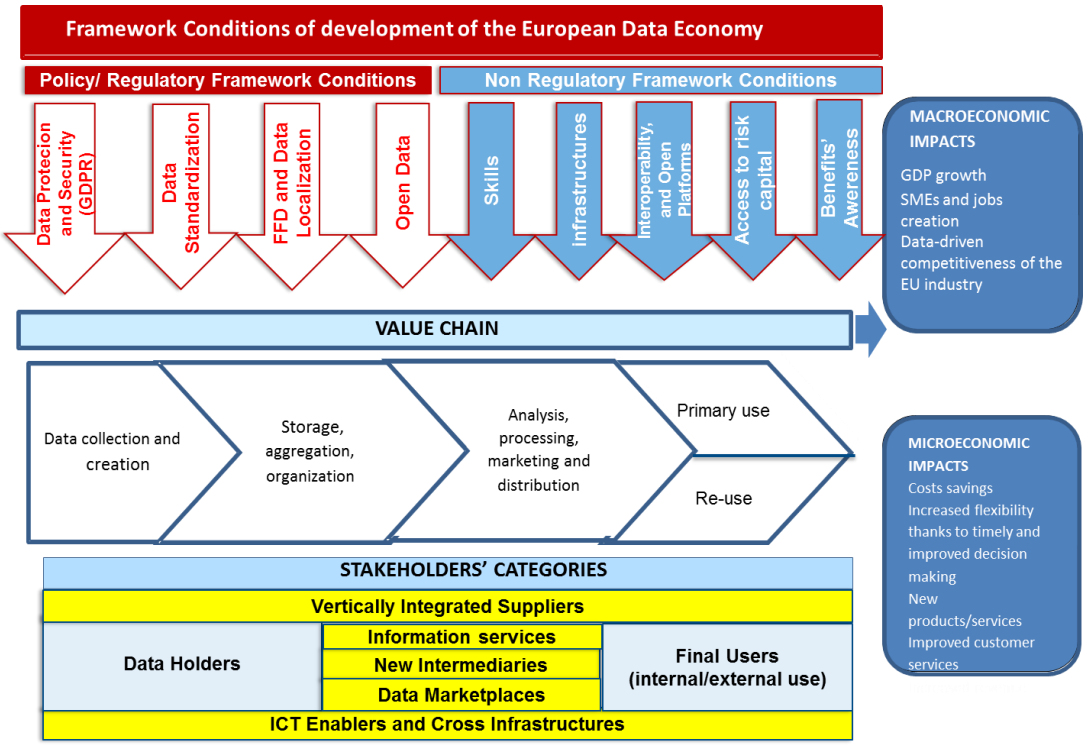
The Data Economy ecosystem is illustrated in Figure 2 below where:

- The value chain shows the four main phases of manipulation of data which lead to its exploitation.
- The macroeconomic and microeconomic impacts identify the direct and indirect impacts of the Data Economy ecosystem on the economic system and user enterprises.
- The stakeholder categories identify the main type of players on the basis of their role in the Data Economy ecosystem.

The main steps in the data value chain are as follows:

- **Collection/access of data** from a myriad of sources within the applicable legal framework. Collection can be direct (for example, through loyalty schemes operated by retailers, transport, and hospitality service providers) or indirect (for example, by recording the location of someone using a cellular phone). Data can be also created through analysis rather than being captured.
- **Storage and aggregation** by service providers and social networks, but also by companies in traditional sectors such as Finance, Retail, Transport, Utilities, and Government.
- **Analysis, processing, marketing, and distribution**, merging data from different sources (public, proprietary, or institutional research) and relying on analytics to derive insights and value. Traditional players across vertical markets can perform this task if they have the necessary skills/technology; alternatively, they can rely on external data brokers and providers.
- **Usage**, both in the public and private sectors, to better serve customers and/or improve efficiency. The use of data is broken down between primary use (when data is used for the goal for which it is collected: for example, mobile traffic data to bill customers by a telecom company) and secondary use or reuse (when data is exploited for other goals, for example when mobile traffic data is used to map customer movements for a retail company). Reuse is expected to be the source of much of the value added in the Data Market.

*Figure 2: The Updated Data Economy Ecosystem*




Source: IDC, 2017

In addition, the data value chain includes the:

- **Framework conditions.** The framework conditions identify the main factors which will enable or prevent the development of the European Data Market and Economy. As indicated in our Data Value Chain design (Figure 2), we have divided the framework conditions into two main groups:
    - Policy/regulatory;
    - Market development/non-regulatory;

## 2.3 The Emergence of a new Data Ecosystem

The Data Market ecosystem and its interrelationships across the data value chain can be apprehended in a more dynamic way. New realities like Data Marketplaces or industrial data spaces should be brought into the mix in order to provide a truthful description of the still-emerging and quickly evolving phenomenon of the Data Market. Data Marketplaces can be defined as third party, cloud-based software platforms providing Internet access to a disparate set of external data sources for use in IT systems by business, government or non-profit organizations. The marketplace operator will manage payment mechanisms to reimburse each dataset owner/provider for data use, as necessary. Along the same lines, but on a more industry-specific level, industrial data spaces are virtual environments facilitating the exchange and connection of data between different organisations through a shared reference architecture and common governance rules. By linking different actors who are interested in sharing information in the form of data, both Data Marketplaces and industrial data spaces constitute a composite business ecosystem combining players from disparate backgrounds, thus fostering the creation of new data-driven services and the emergence of an innovative data ecosystem.

# 3. FORECASTING THE INDICATORS TO 2025

## 3.1 Three Scenarios for the Evolution of the Data Market

### 3.1.1 Overview

The objective of this chapter is to present three potential scenarios of evolution of the European Data Market and Economy to 2025, based on alternative development paths driven by different macroeconomic and framework conditions. These scenarios take as a reference point the 2020 scenarios presented in the previous edition of this study in February 2017[6]. The methodology approach is the same, but of course the key assumptions of the 2020 scenarios have been revised and redefined for the 2025 scenarios, drawing from the following main sources:

- The workshop "Forecasting the European Data Economy to 2025" held by IDC in collaboration with the BDVe CSA project on October 20th, 2017, at the BDVA's premises in Brussels, with the participation of a group of 20 industry and research Big Data experts plus 7 IDC analysts involved with this study. The workshop discussed the main drivers and barriers of the Data Economy and developed several potential scenarios;
- IDC's most recent worldwide Market Forecast Assumptions, dated November 2017;
- IDC's most recent research on Big Data and Analytics, Digital Transformation, Innovation accelerators, and 2018 Predictions on market trends;
- Updated forecasts to 2025 of EU GDP and ICT spending, under 3 alternative scenarios, leveraging the EIU and the IMF data;
- Trends emerging from the EDM Monitoring Tool indicators in the period 2013-2017.

The three scenarios provide the storylines, the contextual framework and the main assumptions which have been used to model and forecast the EDM indicators, with a specific focus on the role of policies. Therefore, the scenarios and the forecast models were developed in parallel, testing their relative coherence and fine-tuning their results. The quantitative scenarios were developed as follows:

- A Baseline scenario was developed first, with the main assumptions based on the continuation of current growth trends and evolution of current framework conditions;

Then we explored the potential alternative development trajectories, resulting in two additional scenarios:

- A High Growth scenario, where the Data Market enters a faster growth trajectory, thanks to more favourable framework conditions;
- A Challenge Scenario, where the Data Market grows more slowly than in the Baseline scenario, because of less favourable framework conditions and a less positive macroeconomic context.

Given the medium-long forecast period, the 2025 scenarios present a wider potential variation of social and business dynamics.

Scenarios are not predictions but potential development paths: their value added lies especially in thinking through the potential consequences of different market trajectories and therefore providing a guide to action. These market scenarios help us to identify the combination of factors and policies best suited to accelerate the development of the data-driven economy in Europe and

[6] "The European Data Market", Final report, February 2017, SMART 2013/0063, http://datalandscape.eu/study-reports

the size of potential economic gains; and conversely to identify the main challenges to this potential growth.

### 3.1.2 Scenario Model and Identification of Key Factors

The scenario model used in this study is based on the definition of alternative assumptions about four main groups of key factors driving the Data Market along different development paths. The identification of the key factors of market development was based on the desk and field research carried out in this study and on the review of a long list of forecast assumptions, leveraging IDC's periodically updated Market Forecast Assumptions. The selection of the most relevant factors was based on two main criteria:

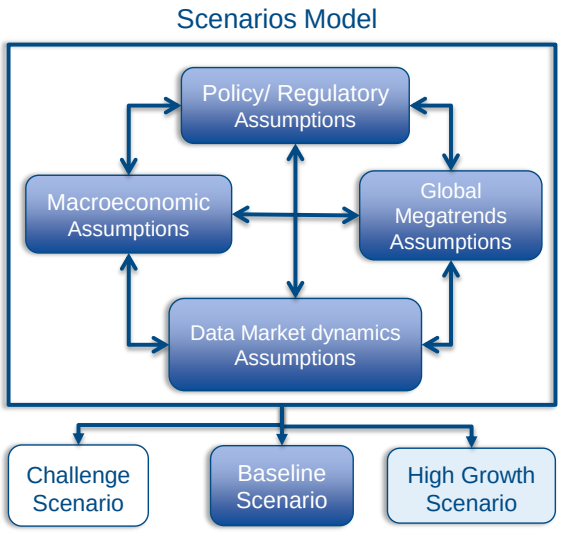
- High level of impact on the development of the Data Market and the Data Economy;
- High level of uncertainty, with potential different outcomes (assumptions) over the next 8 years.

The four main groups of factors are:

- Macroeconomic factors;
- Policy/regulator factors;
- Data Market demand-supply factors;
- Global megatrends affecting all technology markets.

Even though they may seem obvious, 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 3). 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.

*Figure 3: Structure of the Scenarios Model*




Source: European Data Market Monitoring Tool, IDC 2015

The scenario model and the forecast indicators models are correlated. Table 2 below summarises the rationale of their selection and how their assumptions were used as inputs to the indicators' forecast models.

*Table 2: Identification of Main Factors driving the Scenarios*

| Key Factors | Rationale | Inputs to the Forecast Models |
|---|---|---|
| Macroeconomic factors | Macroeconomic factors are partially exogenous to the Data Market (even though data innovation is expected to contribute to GDP growth) | Alternative forecasts of:<br>EU GDP growth 2014-2015-2016-2017-2020 and 2017-2025<br>ICT spending growth 2014-2015-2016-2017-2020 and 2017-2025<br>Other economic factors such as unemployment |
| Policy/Regulatory factors | Policy measures and regulation shape the framework conditions of the development of the Data Market | Alternative policy and regulatory factors by scenario |
| Data Market supply-demand 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 current and forecast ICT innovation drivers as well as global digital transformation dynamics |

The scenarios provide the main framework for the forecast of the EDM indicators. As shown in the Figure below, IDC developed seven forecast models: each model produced the specific indicators forecasts under the three main scenarios, followed by in depth cross-check and quality check. The forecasts models are also correlated and were developed with the following process, with the following dependencies:

- The Data Market forecast model is the cornerstone of the process: it was developed first, building on IDC's forecasts and on the macroeconomic variables as described below. Its results and growth rates feed into the other models, according to the specific assumption and calculation methods explained for each indicator.
- The Data Market and data suppliers'/data users' forecasts influence the data professionals model.
- The data companies' forecasts feed into the data revenues model.
- The data professionals model feeds into the data professionals' skills gap model.
- The Data Economy model feeds from all the other forecasts, but especially the Data Market and the data users' forecasts.

### 3.1.3 Macroeconomic Factors: Background and Relevance

**Background and Relevance for the European Data Market**

The macroeconomic context will have a strong direct impact on the pace of development of the Data Market, influencing the availability of risk capital, the amount of investments, and the willingness to spend on new products and services. We expect the Data Market in the next eight years to be very sensitive to the pace of economic growth, because of the need for ongoing investments and investors' confidence. Another important factor is the pace of growth of ICT spending, which is correlated to GDP. The diffusion of innovative data technologies is positively correlated with overall ICT investments, which include complementary technologies both traditional (servers, network infrastructures) and innovative (cloud computing, mobile and social technologies). However, the pace of growth of the Data Market is faster than the growth of the overall ICT market, which includes a large share of spending for traditional, mature technologies. According to IDC's forecasts, investments in data technologies will continue to outpace traditional ICT investments in the next eight years, as enterprises develop their data supply chains and their data-related products and services. As the Data Market grows in size and data-driven innovation deploys its benefits, its

contribution to economic growth will become more visible generating a virtuous cycle of development. In addition to the qualitative assumptions, we have developed alternative estimates of GDP growth and ICT investments under the three main scenarios, for the EU and each of the EU28 Member States. We used these variables to cross-check and fine-tune the forecast estimates of the Data Market growth. The main data are presented in the tables below. By 2025 the U.K. should be out of the EU so we present both EU27 and EU28 data.

- The forecast EU27 GDP cumulative growth rate for the period 2017-2025 ranges between 0.7% for the Challenge scenario, 1.3% for the Baseline and 2.2% for the High Growth scenario. The EU28 growth rates are slightly lower because of the U.K. economy lower growth forecasts due to Brexit.
- The forecasts of ICT spending for the Baseline scenarios are sourced from IDC's Black Book[7] data base, 4th Quarter 2017. The Challenge and High Growth scenarios' estimates were done leveraging IDC's database historical series of ICT to GDP correlations. EU27 ICT spending growth is slightly lower than the EU28 because the U.K.'s ICT spending (one of the largest in Europe) is expected to grow faster than the other MS. However, total ICT spending will grow much more slowly than the Data Market and GDP because of the decline in traditional IT spending.
- The combination of these trends results in a relatively stable share of ICT spending on GDP in the 3 scenarios, higher in the Baseline scenario (4.1%). The share of the Data Market on ICT spending, instead, is predicted to raise quickly in all 3 scenarios but naturally much faster in the High Growth one. The U.K. forecasts push EU27 ratios slightly higher because we do not expect Brexit to slow down by much the U.K. Data Market, which is very strong.

*Table 3: Macroeconomic Variables used for the Forecast Model (€, Billion)*

| | 2016 | 2017 | 2020 Baseline | 2025 Challenge | 2025 Baseline | 2025 High Growth |
|---|---|---|---|---|---|---|
| ICT Spending | | | | | | |
| EU27 | 472 | 483 | 508 | 519 | 536 | 553 |
| EU28 | 608 | 624 | 657 | 679 | 702 | 724 |
| GDP | | | | | | |
| EU27 | 11,744 | 12,015 | 12,702 | 13,175 | 13,551 | 14,148 |
| EU28 | 13,821 | 14,128 | 14,918 | 15,429 | 15,923 | 16,653 |
| Data Market | | | | | | |
| EU27 | 46 | 50 | 60 | 74 | 85 | 110 |
| EU28 | 59 | 65 | 77 | 96 | 110 | 146 |

Source: European Data Market Monitoring Tool, IDC January 2018

*Table 4: Macroeconomic Variables used for the Forecast Model – Ratios (%)*

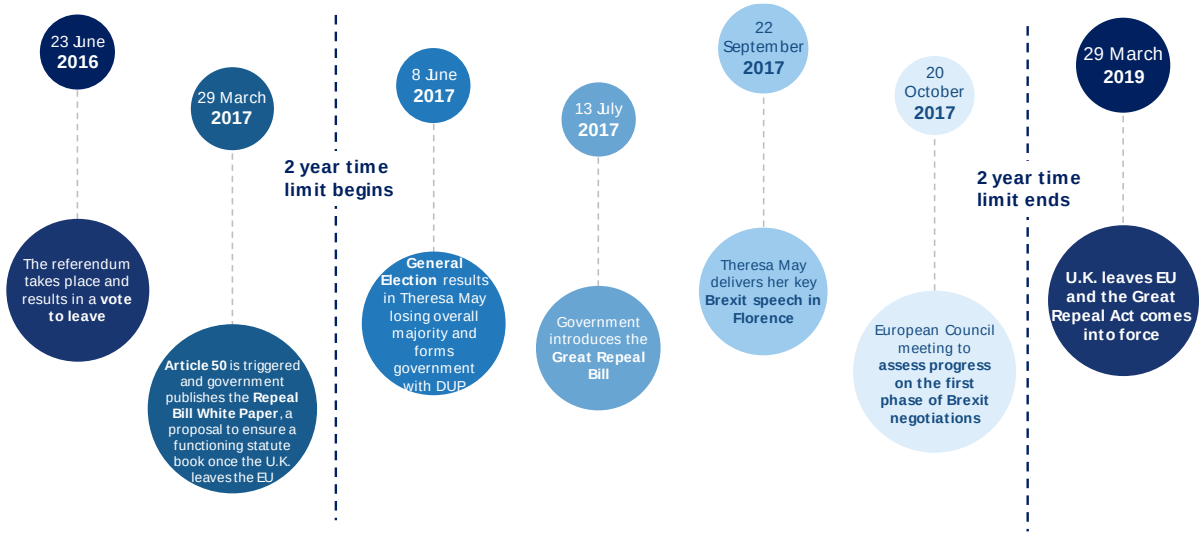
| | 2016 | 2017 | 2020 Baseline | 2025 Challenge | 2025 Baseline | 2025 High Growth |
|---|---|---|---|---|---|---|
| ICT Spending/ GDP | | | | | | |
| EU27 | 4.0% | 4.0% | 4.0% | 3.8% | 4.1% | 3.9% |
| EU28 | 4.4% | 4.4% | 4.4% | 4.3% | 4.5% | 4.3% |
| DM/ ICT spending | | | | | | |
| EU27 | 9.8% | 10.4% | 11.9% | 14.2% | 15.9% | 19.8% |
| EU28 | 9.8% | 10.4% | 11.8% | 14.2% | 15.7% | 20.2% |

[7] 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

### 3.1.4 Brexit Perspectives

The common opinion is that the consequences of Brexit are even more uncertain now than at the date of the U.K. referendum of 23 June 2016 to leave the European Union. The timeline of events is presented in the figure below. According to Premier Theresa May announcement, the U.K. will definitely leave the EU on 29 March 2019 and the Great Repeal Act will come into force, but a further period of transition is likely before final separation is enacted.

*Figure 4: Brexit Timeline*




Source: IDC November 2017

IDC analysed the Brexit situation and its impact on the IT market in November 2017[8], with the following considerations.

- U.K. GDP growth decreased across 2016 and 2017. Despite a small increase from 2Q17 to 3Q17, this has not prevented the U.K. from registering the lowest annual growth rate since 2012, which is particularly poor considering that the global economy is benefiting from a cyclical upturn. Moreover, consumers reduced their spending in the light of increasing inflation.
- Uncertainty remains and affects expected GDP growth. The plan to eliminate the deficit by the mid-2020s will be harder if growth is lower.
- U.K. inflation reached a peak at the beginning of 2017. It is slowly decreasing but remains very high compared to most other European countries. Inflation is affecting almost all sectors of the economy.
- The fall in the pound since last year's Brexit vote has been one factor behind the rise in the inflation rate, as the cost of imported goods has risen, prices and a range of transport prices in particular have been hit. As inflation is now 0.9% above the rate of wage growth, consumers face real-term declines in household income.
- The rise in inflation pushed the Bank of England to its first-rate hike since 2007, increasing interest rates from 0.25% to 0.5%, with the effect of increasing the price of mortgage repayments for households.
- European Services and Software Market forecast is quite stable when compared with the first half of 2016, a trend confirmed and slightly accentuated in the U.K. market. IDC has not

[8] IDC European IT Market Update – A deep dive on the U.K., https://www.idc.com/getdoc.jsp?containerId=EMEA43282117

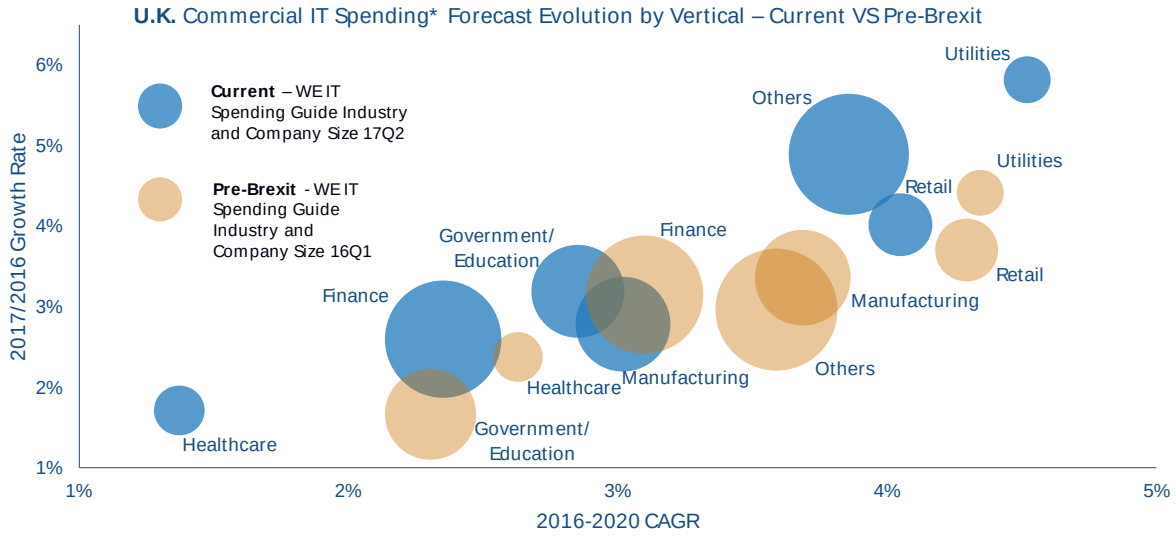
seen any disruptive Brexit impacts on these two markets yet, which are driven by multiyear outsourcing, maintenance, and enhancement contracts.

- The Hardware market, on the other hand, drives the overall IT market forecast change. An increase in the average price of devices, driven by both currency fluctuation and price increase, boosted hardware market growth, despite a unit drop and uncertainty in some subsectors and countries. The U.K. hardware market shows weak positive expected growth for the forecast period.
- Public sector, Finance, and Manufacturing are the most exposed British industries to Brexit effects and ongoing negotiations. This is driven by large organisations' relocation moves and currency fluctuation in the shorter term, and the recruitment gap, new internal rules and relationships with other countries in the longer term.
- Impacts will be more limited for Utilities and other industries, such as Telecoms/Media, Transport, Professional services, and Personal & Consumer services.

In summary, uncertainty is affecting the perspectives of growth of the U.K. economy and IT industry through the currency rate and inflation, even though major negative impacts are not yet visible. The type of trade agreement to be reached with the EU is still completely uncertain and the option of a Hard Brexit is not off the table. Major corporations are keeping their options open and slowly moving European headquarters and jobs outside of the U.K.

In this context, IDC has dampened down the growth perspectives of the U.K. Data Market and Data Economy in all 3 scenarios, but not in a dramatic way. Brexit impacts are likely to evolve very slowly over a period of several years. The drivers of the Data Market are very strong in a service economy like the U.K. with a high incidence of advanced data users, so even if the economic conditions will worsen the Data Market is still likely to grow. However, the degree of uncertainty about the U.K. is so high that we advise to consider separately the EU27 data from the U.K. data.

*Figure 5: Forecast Evolution of U.K. IT Spending by Vertical Market, 2016-2020*




Source: IDC November 2017

### 3.1.5 Policy/Regulatory Conditions: Background and Relevance

The potential role of policies in meeting the main challenges of development of the emerging Data Market and Data Economy has been discussed in depth over the past few years, within and outside the EC. This paragraph focuses on the identification of the policy initiatives with the most direct impact on our forecast scenarios, and a high degree of uncertainty, leading to different policy

assumptions for each scenario. To do so, we have focused on the EC-driven policies, as they cover all the main policy issues relevant for the Data Market and inspire national governments actions.

In the EC context, different strands of policy elaborated in the last years about data-related challenges (including the Open data and the Data Value Chain policy) were brought together and systematized in the EC Communication "Towards a thriving data-driven economy" (July 2014), defining a coherent Action plan, broadly consistent with the OECD analysis but with different priorities. The Action Plan outlines also the interactions and synergies with other policies and initiatives, such as the Horizon 2020 (H2020) Programme for collaborative research and innovation and the Digital Entrepreneurship policy. The Digital Single Market strategy (DSM) launched in May 2015 confirmed the relevance of this issue, targeting the growth of the data-driven economy as one of its key priorities. Some of the actions of the DSM Roadmap are designed to pursue key policy objectives anticipated by the data-driven economy strategy.

In the year 2016, a new set of initiatives were launched by the EC to further sustain the data-driven economy in Europe. In April 2016 the EC published a new Communication on the digitization of the European industry aimed at improving the digitisation process across several economic sectors in Europe[9]. The digitisation of the European Industry would serve as coordination platform for other European, national & regional initiatives (such as Industrie 4.0 in Germany, Smart Industry in the Netherlands or the Nouvelle France Industrielle in France) and is accompanied by other initiatives in the field of standards[10] (to promote widely accepted standards in priority areas such as Big Data, 5G, Cloud Computing, Internet of Things and Cybersecurity), accelerating the digitisation of public services[11] through a rejuvenated eGovernment Action Plan and reinforcing the uptake and strengthen the benefits to be derived by Cloud technologies[12] in a data-driven economy. At term, all these initiatives will produce results that need to be measured at regular intervals to gauge their progress towards their specific objectives and take corrective actions if necessary.

Other important initiatives were launched in 2017. The EC Communication "Building a European Data Economy" published in January 2017 outlined the main priorities of a policy framework for the Data Economy, addressing the following issues: removing barriers to the free flow of data across borders; ensuring data access and transfer; dealing with new liability and safety issues for emerging technologies such as automated vehicles and robots; enabling the portability of non-personal data, interoperability and standards. A major step to act on these priorities is the proposed Regulation framework for the Free flow of non-personal data, published in September 2017, which should be finalised by the end of 2018 after the negotiation with the Council of Ministers and the European Parliament.

Again, in January the EC proposed an update to the e-privacy regulation. The package included new data protection rules for EU institutions and set out the Commission's strategic approach to the issues concerning international transfers of personal data. Finally, a political agreement was reached

[9] Digitising the European Industry – COM (2016) 180 final

[10] ICT Standardisation Priorities for the Digital Single Market – COM (2016) 176 final

[11] EU eGovernment Action Plan 2016-2020, Accelerating the digital transformation of government – COM (2016) 179 final

[12] European Cloud Initiatives – COM (2016) 178 final

in November 2017 between the European Parliament, the Council and the Commission to end unjustified geoblocking, an item on the agenda since 2016.

The finalisation of these initiatives is included in the EC legislative plan for 2018-19. The Commission has set itself the objective to wrap up legal negotiations on all 25 of the Digital Single Market proposals that it announced since 2015 (of which 13 are still open). The most relevant for the European Data Market are indicated below. The Commission spokeswoman Nathalie Vandystadt named ePrivacy, cybersecurity, copyright and the free flow of data as "cornerstones" of the now nearly three-year-old Digital Single Market agenda. Their completion and eventual implementation will shape the framework conditions for the European Data Market.

*Table 5: EC Legislative Priorities 2018 – Digital Single Market*

| EC Legislative Priorities 2018-19 – DIGITAL SINGLE MARKET | Level of Impact on the Data Market |
|---|---|
| Proposal for a DIRECTIVE establishing the European Electronic Communications Code | Low |
| Proposal for a REGULATION establishing the Body of European Regulators for Electronic Communications | Low |
| Proposal for a REGULATION for the free flow of non-personal data in the European Union | High |
| Proposal for a DIRECTIVE on certain aspects concerning contracts for the supply of digital content | High |
| Proposal for a DIRECTIVE on certain aspects concerning contracts for the online and other distance sales of goods | High |
| Proposal on fairness in platform-to-business relations | High |
| Proposal for a DIRECTIVE on copyright in the Digital Single Market | Medium |
| Proposal for a REGULATION concerning the respect for private life and the protection of personal data in electronic communications | High |
| Proposal for a REGULATION on the protection of individuals with regard to the processing of personal data by the Union institutions, bodies, offices and agencies | High |

*Table 6: Recent EU Initiatives and Level of Impact/Uncertainty at 2025*

| Additional Initiative | Level of Impact on the Data Market Development | Level of Uncertainty of Potential Outcomes, 2025 |
|---|---|---|
| Digitising the European Industry | High | High: difficulties in coordinating existing national- and regional-level initiatives and while obtaining buy-in and commitment from the private sector |
| ICT Standardization Priorities for the DSM | High | High: uncertain success of the policies in terms of adoption and take-up of guidelines by all MS |
| Digitalising Public Services | High | High: existing disparities in digitisation capabilities and ICT skills among Public Administrations across the EU |
| European Cloud Initiative | High | High: difficulties in obtaining wider access and building trusts among economic operators in Europe |

Source: European Commission 2016 and IDC elaboration

In summary, the following table presents the main policy and regulatory conditions with both a high impact and high level of uncertainty of potential outcomes by 2025. They have been used to develop the relative policy and regulatory assumptions for the alternative scenarios.

*Table 7: Summary of Policy and Regulatory Conditions used to develop Scenario Assumptions*

| Policy/Regulatory Conditions | Policy and Initiatives | Level of Uncertainty |
|---|---|---|
| **Personal Data Protection and Consumer Protection** | GDPR, e-Privacy Directive | Uncertainty about the effectiveness of implementation and the need of adaptation by enterprises |
| **Promote portability of non-personal data, interoperability and standards** | Free Flow of Data Proposed Regulation | Approval likely, but not certain, within the mandate of the current Commission<br>If approved, positive impacts depend on MS removal of data location restrictions and industry complying with new codes of conduct |
| **Promote the Free-flow of data and improve access and transfer of data** | Free Flow of Data Proposed Regulation | |
| **Fostering Open Data policies** | Public Consultation and proposed revision of the PSI Directive 2013/37EU (2018)<br>Open Research Data Policy<br>European Data Portal funded by the Connecting Europe Facility (CEF) | Even though open data is growing, there are still differences across the EU about development and exploitation |
| **Developing a skills base for the** Data Economy | Digital Skills and Jobs Coalition initiatives<br>Data skills development by the European Big Data Centres of excellence | Uncertainty about the effectiveness and timing of development of the data skills needed |

### 3.1.6 Data Market Dynamics' Factors: Background and Relevance

The pace of growth of the European Data Market in the next years will be strongly influenced by Europe's capability to develop a healthy supply-demand ecosystem. This paragraph focuses on the key factors which may lead to alternative supply-demand dynamics correlated with faster or slower Data Market development paths.

As with most innovative markets, the Data Market was launched by the push of innovative technology offerings (technology push), but will reach its full potential only when demand is sufficiently mature (demand pull). The analysis of the current Data Market in Europe, presented in this report, shows a dynamic supply-side (an emerging data industry) and a not-yet fully developed demand-side, where actual users are still a minority of potential users. Technology-push is still the dominant model of supply-demand interaction.

The 2020 scenarios were focused on the potential transition from a supply-push model in the Baseline scenario to a demand-pull model in the High Growth scenario. Moving the horizon to 2025 the potential evolution of the Data Market becomes more complex and more factors come into play, as outlined below in the following tables describing the main drivers and barriers considered by the scenarios.

*Table 8: Data Market Drivers*

| Supply-Demand Drivers | Impact | Likelihood | Rationale |
|---|---|---|---|
| Globalisation and international competition/cooperation | H | H | Organisations will need to adopt data-driven innovation and digital transformation to keep up with international competition and global supply chains |
| Diffusion of data-driven business models and increase of data-driven organizations<br>Ability to monetize data assets | H | H | More and more organisations will adopt data-centric models, engage in digital transformation, experiment with new business models and offer data services |
| Increased consumer demand of data-based services to manage frantic lifestyles | H | H | Consumers will need apps and devices to be always in touch, manage family, work and leisure |
| Servitisation of industry increases business demand of data services | H | H | The transformation of industrial products into services relies on data |
| Need to improve efficiency/productivity/reduce costs | H | H | International competition drives increasing pressure on costs - digital processes enabled by data improve productivity and efficiency |
| Availability of / access to new data sources, data flows (also cross-borders), data sharing | H | H | Successful implementation of the Digital Single Market will enable enterprises to grow faster and achieve economies of scale |
| Diffusion of Digital Innovation Hubs accelerating innovation | H | H/M | Regional centres of excellence will develop data skills, provide space for enterprises experimentation and learning |
| Proving the business case for data sharing (also for SMEs) | H | M | The emergence of multi-user digital platforms and pioneer experiences will demonstrate the viability and profitability of data sharing business models |
| Increased maturity/competitiveness of supply also thanks to standardization | M/H | M | European industry will improve its competitiveness building multi-user digital platforms and leveraging standards |
| Decreasing costs of data technology | M | H | As the market matures solutions and tools will become more affordable for more enterprises |
| Lower barriers to new markets for SMEs | M | M | SMEs learn to enter new markets thanks to data-driven innovation |
| Emergence of circular economy drives demand of data | M | M/L | Environmental policies require the efficient management of products lifecycles and waste, based on pervasive business intelligence and innovative data services |

Source: IDC elaboration on Data Economy scenario workshop, 2018

*Table 9: Data Market Barriers*

| Barriers | Impact | Likelihood | Rationale |
|---|---|---|---|
| Enterprises uncertainty in the implementation of GDPR | H | H | Great uncertainty about practical implementations may lead to missed business opportunities and in time to GDPR as constraint. For example: uncertainty about treating IoT data |
| Business fear to lose power/ control of own data | H | H | IPR issues and concern of large enterprises taking advantage of SMEs major barrier preventing data sharing |
| Lack of trust in data suppliers | H | H | Many enterprises do not trust providing their data to data suppliers - also EU companies tend to mistrust U.S.-owned leading platforms |
| Lack of awareness of data sharing benefits and business opportunities | H | M | Scarcity of proven business cases and lack of awareness/ knowledge of existing or potential benefits, particularly for SMEs, particularly of data assets monetization |
| Lack of data strategies/ capability to extract-exploit data | H | L | Only large organisations have data strategies but in time most will understand the need to improve their capability to make their data usable and monetize it |
| Uncertainty about the TCO of data (total cost of ownership) | M | H | Again, the cost-benefit balance of data investments is not well understood - and still needs to be documented |

| Barriers | Impact | Likelihood | Rationale |
|---|---|---|---|
| | | | particularly for IoT and AI services |
| Increased market concentration/ dominance by few mostly non-EU competitors | M | H | The winner takes all and network effects concentrate power in few global companies which may impose business conditions and pricing to all other enterprises (GAFA: Google, Apple, Facebook, Amazon). Need for updated competition and antitrust regulation |
| Protectionism and vested interests slow down disruptive data technology innovation | M | M | The incumbents in traditional industries such as finance and manufacturing may block disruptive innovation with the help of governments |
| Europe's diversity increases costs/reduces economies of scale | M | M | Fragmented markets, variety of languages, slow down growth opportunities for European innovative data companies both in terms of availability of data and of reaching a sufficient number of customers |
| Insufficient availability of risk/venture capital in Europe | L | H | Risk capital has been historically low in Europe, availability has improved for greenfield and early start-up phase, still difficult to find development capital for EU companies unless they agree to relocate to U.S. |
| Lack of digital skills by data users | L/M | L | Even if with low probability, still a lack of data skills in user businesses may prevent data-driven innovation rapid diffusion and adoption in traditional and smaller enterprises |

Source: IDC elaboration on Data Economy scenario workshop, 2018

### 3.1.7 The Social Role of Data

The new wave of innovation driven by IoT, Artificial Intelligence and Robotics and the pervasiveness of the role of data in our personal and working life are creating new ethical and social challenges and re-shaping social interaction. The increasing dominance of a few global platforms (GAFA – Google, Apple, Facebook, Amazon) controlling a massive amount of data not only poses new challenges for competition and antitrust legislation, it creates new worries for equality, social cohesion and democracy (such as the difficulty to manage fake news and the risks of manipulation of political competition). For this reason, we have decided to pay greater attention in the 2025 scenarios to social factors and the social role of data. A summary of the main social drivers and barriers that will influence the Data Economy in the next years and may accelerate or slow down the diffusion of data-driven innovation is presented below.

*Table 10: Social Drivers*

| Social Drivers | Impact | Likelihood | Rationale |
|---|---|---|---|
| **Incentives/ compensation/rewards to exploit own data** | H | H | For example, pay-with-your-data type of Business Models will develop and diversify |
| **Improvement of social and/or physical wellbeing** | H | M | Diffusion of fitness and wellbeing culture with high social value |
| **Increase of participation and/or more participatory democracy** | H | M | Social media potentially enabling greater participation but also risks of manipulation of public opinion (fake news, populism) |
| **Growth of data sharing culture** | M | H | Social evolution: greater willingness to share personal data, different perception of privacy, people becoming open to data sharing also in business environment. |
| **Momentum and peer pressure** | M | M | Momentum drives success of social platforms (Facebook) but may change with platform migration (even though "winner takes all" and network effects sustain leaders) |

Source: IDC elaboration on Data Economy scenario workshop, 2018

*Table 11: Social Barriers*

| Barriers | Impact | Likelihood | Rationale |
|---|---|---|---|
| Lack of trust in data sharing in personal and business virtual environments | H | H | Distrust comes from the gap between knowledge of the world and acting in it (reasonable expectation test is untestable) and from increasing awareness of lack of reliability of content on the Internet (fake news, exc.) |
| Fear of external control / loss of self-determination because of data abuse | H | M/H | Policy-cycles and society in general can put barriers due to dystopian views entering the political decision-making process |
| Increase of cybercrime and/or fear of cybercrime | H | H | Increase of data dependency will also lead to an increase in data breaches and profitability of cybercrime (no technology solution will be able to completely prevent cybercrime, not even blockchain) |
| Insufficient digital literacy | H | H | Knowledge gaps continue to increase, despite coding classes at school - inherent complexity also makes for inherently different types of knowledge |
| Prevalence of winner-takes-all business models | H | H | Leading to concentrated markets where the winner has all the power and aggressive business strategies have great rewards. Risk for EU innovative companies who are less likely to be winners and may become prey for acquisitions. GDPR may become constraint for EU companies |
| Impact of cultural differences slowing down data-driven innovation | H | M | Fragmented markets, language barriers weaken the diffusion of data-driven innovation; harmonisation and open standards may create a common ground (although also to digital colonialism) |
| Resistance to social change and technical innovation | H | M/H | There could be an increase or influx of part of society that opt out and want out - especially in an ageing society - government has to slow down and adopt multiple speeds of innovation implementation |
| Pessimistic vision of the future | M | H | Driven by distrust and fear of poorly understood new technologies, may undermine willingness to accept innovation |
| Increased social control, potential censorship and reduction of freedom | M | M/H | Chilling effects due to ubiquitous tracking, both on the online-and offline world, under the banner of rationalisation and efficiency, more and more spaces will be tracked, increasing people's fear to behave as they wish |
| Lack of data sharing culture | M | H | Lack of trust, fear of cybercrime, social resistance to innovation may combine to slow down data sharing |
| Increased digital divide | M/L | M | ICT pervasiveness in all social and work interaction may worsen the digital divide unless this is actively prevented |

Source: IDC elaboration on Data Economy scenario workshop, 2018

### 3.1.8 Global Megatrends: Background and Relevance

Digital innovation is driven by global trends affecting the whole world. The combination of Big Data, Cloud Computing, Mobile technologies and Social media is the most powerful driver of change of the economy and the best opportunity for Europe to move back to a growth path. In addition, innovation will be accelerated in the next years by new developments such as Cognitive Systems, Robotics, 3D Printing and most definitely by the IoT, the Internet of Things, whose networks of sensors will generate huge amounts of data and create "smart environments". The interconnection of these technologies is spreading to all industry sectors, pervading and reshaping business processes and leading to the digital transformation of all enterprises, without exception. In the last year however, new technology trends are arising which will be very important for the Data Market: they include for example blockchain, 5G infrastructures and PET (Privacy Enhancing Technologies). The table below lists the many new technologies which will need to be implemented for Europe to keep up pace with global macrotrends and the Data Economy to achieve the best possible growth.

*Table 12: Global Technology Drivers*

| Drivers | Impact | Likelihood | Rationale |
|---|---|---|---|
| Diffusion of AI / Cognitive computing | H | H | Extremely important but speed of technology evolution and adoption still uncertain |
| Diffusion of IoT/Sensors everywhere/wearables | H | H | Diffusion and take-up; continuing trend, already happening and even more throughout 2025 |
| Diffusion of Cloud computing/fog-edge computing | H | H | Principle way of data storage and transfer - fundamental technology for a Data Economy, already established and even more prominent in 2025 |
| Diffusion of business-oriented digital platforms | H | H | Clear growth trend, even though issues remain concerning data sharing and data ownership |
| Improved Security technologies/improving trust | H | H | Same as above, especially in the light of the take up of distributed ledger technologies and blockchain |
| Diffusion of Digital twins of products and services | H | H | Development of digital copies relies on data, enables new services, data-driven innovation |
| Diffusion of New medical sensing technology | H | H | Also a sub-category of IoT endowed with a multitude of sensor and actuators - specific use case of IoT, similar trend as above |
| Diffusion of blockchain | H | H | Huge potential and on the rise but many uncertainties related to trust and potential resistance by vested interests (incumbent finance and insurance operators). Currently applied in a relatively limited number of use cases (payments, transfer of data for smart contracts, etc...) |
| Diffusion of Technology for data provenance (traceability) for healthcare, pharmaceuticals, food-agriculture | H | H | Traceability technologies other than blockchain will acquire importance over the next few years in several use cases |
| Diffusion of PET (Privacy Enhancing Technologies) | H | H | As privacy and trust remain key in the data sharing debate, PET are expected to play a prominent role in 2025 |
| Diffusion of 5G, HPC infrastructures | M | H | 5G networks and devices should start being deployed around 2018-19 and commercial services from 2020 onwards (source: IDC), but the rate of development of standards is uncertain |

Source: IDC elaboration on Data Economy scenario workshop, 2018

*Table 13: Global Technology Barriers*

| Barriers | Impact | Likelihood | Rationale |
|---|---|---|---|
| Complexity of data slowing down exploitation | H | H | The usability of data is conditioned by too many parameters, too many sources, functionalities, levels, terminology, interaction with legacy systems |
| Lack of standards, insufficient interoperability | H | H | Risk to slow down innovation, particularly IoT diffusion |
| Lack of connectivity/insufficient deployment of infrastructures/ slow deployment of 5G slowing IoT | H | H | Risk of lack of coverage in some areas particularly rural areas, creating patchwork demand |
| Blockchain shortcomings (technical immaturity, lack of trust or reliability) | H | M | Uncertainty about scalability (what if chains get to trillion logs?) security interoperability: solutions needed |
| AI-driven algorithms reliability and capacity to work together still a challenge | H | M | The promise of new services working smoothly together is not fulfilled and applications remain siloed |
| Robotics/ ethical, liability and accountability issues | M | M | Ethical and liability issues still emerging and not solved. Who is responsible if automated cars kill people? Or robots in factories create damages or accidents? |
| Insufficient accessibility and availability of data | L | M | Risk that legal barriers, insufficient data skills constrain the availability of the vast amount of data needed to enable the Data Economy |

Source: IDC elaboration on Data Economy scenario workshop, 2018

## 3.2 Overview of 2025 Scenarios

While the 2020 scenarios were mainly differentiated by economic drivers (different demand-supply dynamics), the 2025 scenarios are 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** (Data innovation Waves) is characterized 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 characterized 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.

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.

## 3.3 Description of the Baseline Scenario

As the world economy enjoys a positive growth period, this scenario predicts a continuation of these positive trends to 2025 with a cumulative GDP average growth rate of 1.3% over the period 2017-2025, which is 50% higher than in the Challenge scenario and 70% lower than in the High Growth. This sustains the investments in the digital economy and consumer willingness to spend. The need to manage dynamic lifestyles always on the move will continue to sustain the demand for data apps and services. Responding to globalisation and international competition, European enterprises will accelerate their digital transformation process, enabled by data-centric processes and new digital

business models. In 2017, according to IDC's maturity benchmark statistics, 55% of European medium and large organisations were still shy of full digital transformation, between the digital explorer and digital player stages. By 2025 we expect a majority of large enterprises over 250 employees and all innovative SMEs to become data-driven organisations, experimenting with new business models. Already, IDC estimates that by 2020, 50% of large enterprises will be generating Data-as-a-Service revenue from the sale of data assets.

The monetization of data assets and the diffusion of digital multi-user, cloud-based platforms will enable B2B data sharing and spread awareness of the potential benefits of data exploitation. Demand-supply dynamics will move towards a mix of continuing technology push with demand pull. Several business trends will push the demand of data services: the industry move towards "servitisation", transforming products into services (for example, selling mobility services instead of cars), the need to keep costs under control, the decreasing costs of data technology, the diffusion of "digital twins" of products and services built through data to simulate/anticipate faults, performances, check quality.

This scenario foresees a healthy growth of the European data industry, a continuing improvement of the offering of data products and services, and a corresponding gradual development of demand, especially by the most advanced, competitive and innovative enterprises, large and small.

Nevertheless, this is a scenario of "patchwork" innovation across Europe, by industry and by region, as relevant barriers still remain. In this scenario, policy will play an important role to support supply, but have a mixed success in promoting demand, an inherently more difficult objective. Policy initiatives will succeed in supporting the growth of the data industry through R&D investments, the support of digital entrepreneurship, and the successful deployment of the contractual Public Private Partnership on Big Data Value (BDVA PPP). The EU will protect trust in the Data Economy by successfully deploying the General Data Protection Regulation, achieving greater harmonization across the EU and reducing the administrative burden on businesses, but uncertainty about implementation and fear of fines may prevent some enterprises from experimenting with innovation. The Free flow of data initiative will have some success to remove data location barriers, but will take time to be fully implemented towards the end of the forecast period.

The leading global players (Google, Facebook, Amazon, Apple) will maintain their dominance of personal data flows, but will have to negotiate allowing individual users some control on their personal data and accepting some transparency measures. Anti-trust and competition authorities across Europe and the Americas will work hard to develop tools and methods able to contain the power of the big platforms and impose fiscal obligations. Open data platforms and industry platforms will dominate non-personal data flows creating a more open and competitive Data Market. This scenario therefore is positioned between the two extremes of high and low concentration of power and data control.

## 3.4 Description of the High Growth Scenario

In the High Growth scenario, Europe's economic growth in the next years will be higher than the Baseline scenario and will be characterised by a stronger driving role of digital innovation, with higher overall ICT investments as a share of GDP. Solutions combining innovative digital technologies (such as IoT, Cloud and Big Data) will be more widely implemented and more European enterprises will engage in Digital Transformation before 2020. The Data Market will enter a faster growth trajectory and the adoption of data technologies will spread beyond the minority of pioneers to a wider population of mainstream users. 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. This is a classic virtuous cycle mechanism, which may happen if data technologies take-up starts climbing fast enough to generate momentum. Because of network effects typical of ICTs, rapid diffusion multiplies the benefits for users in their interactions and makes it easier to consolidate standards and interoperability, reducing further the barriers to adoption.

To enable this scenario, we must assume a set of very favourable framework conditions, which will be able to trigger a faster take-up. First of all, the adoption of all digital technologies is mutually reinforcing, so we assume a faster pace of diffusion for IoT, Cloud, Mobile as well as data technologies. Second, we must assume a leap ahead of awareness of potential benefits and willingness to adopt data technologies by mainstream users and specifically by SMEs. Third, but not less relevant, we must assume a removal of existing regulatory barriers within the forecast period. In this scenario, policy initiatives will succeed in supporting supply as detailed above, but will also have better success in promoting demand. Policies enabling the free flow of data cross-borders and the re-use of data sets will create positive effects on demand already from 2019-2020. All the other positive factors described in the Baseline scenario must also be present.

## 3.5 Description of the Challenge Scenario

In the Challenge scenario, the combination of a less positive macroeconomic context than in the Baseline scenario, less favourable framework conditions, and slower diffusion of digital innovation, will combine to push the Data Market into a low growth development path. This is a fragmented scenario, where the Digital Single Market will fail to materialise before 2020. The supply-demand dynamics will be dominated by the technology push, since the demand pull will be weak. This scenario therefore explores the potential risks and consequences of failing to remove the barriers to the development of the Data Economy in Europe.

This scenario still anticipates an increase of the diffusion of digital technologies such as IoT and Cloud, but at a slower pace than in the Baseline. The dynamics of mobile and social technologies should not be much different in this scenario, given their strong momentum and their closeness to nearly universal diffusion. As a result, the “hyperconnected” society will become closer in this scenario too, even if less well developed than in the Baseline or High Growth scenarios. It is possible that the diffusion of high-speed broadband infrastructures across Europe will be incomplete, with the risk of a digital infrastructures divide between and within the Member States. This will be another element of weakness for the development of the Data Market.

In this scenario, both supply-side policies and demand-side policies will tend to have weaker impacts and to be deployed more slowly in time. Policy initiatives will still succeed in supporting the growth of the data industry through R&D investments, the support of digital entrepreneurship, and the successful deployment of the BDVA PPP, but to a lesser extent given the lower propensity to invest by the private sector. Policies addressing enabling conditions, such as the removal of regulatory barriers to the free flow of cross-border data, will be delayed in time and be less effective than in the Baseline scenario. As a consequence, the value of the Data Market and of the Data Economy by 2025 will be substantially lower than in the Baseline scenario.

# 4. MEASURING THE DATA PROFESSIONALS

## 4.1 Definition

***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. They elaborate and visualise structured and unstructured data to support analysis and decision-making processes.*

As specified in the Inception Report (D1), in this Update of the European Data Market Study, data professionals are not only data technicians but also data users who, based on sophisticated tools, take decisions about their business or activity, after having analysed and interpreted available data. As a result, data professionals belong to the category of knowledge workers and specifically "codified" knowledge workers (Lundavall and Johnson, 1994). More precisely, data professionals deal with data, while knowledge workers deal with information and knowledge.

## 4.2 Measuring Data Professionals

The process for estimating data workers in the previous European Data Market Study (SMART 2013/0063) has been substantially replicated to identify data professionals in this update report. The study team, however, has extended the list of sources and gleaned into an additional set of data to account for a definition of data professionals that includes individuals with the ability to take decisions on key business and/or organisational aspects of their job.

The sources used by the study team include the EU Labour Force Survey (EU LFS), International Labour Organization's ILOSTAT Database, Eurostat's Structural Business Statistics (SBS), and the Eurostat Business Demography. The table below outlines the main sources used to estimate the data professional indicators in this report. The full and updated list of data sources, as well as other methodological details, is presented in the Methodology Annex at the end of this document.

*Table 14: Data Professionals – Main Data Sources*

| Data Source | Updated |
|---|---|
| Eurostat Business Demographic Statistics | Dec 2017 |
| Eurostat Annual Structural Business Statistics | Dec 2017 |
| IDC Worldwide Black Book (standard edition) | Nov 2017 |
| ILOSTAT statistics and databases | Jan 2018 |
| The European Union Labour Force Survey (EU LFS) | Oct 2017 |

All in all, data professionals are estimated at a total of 5.3 million in the EU27 and at 6.7 million in the EU28 in 2017, thus marking a significant increase in 2017 over the previous year (8.2% and 8% year-on-year respectively). When compared to the year 2017, the 2020 Baseline scenario outlined in the previous European Data Market study (SMART 2013/0063) would register a Compound Annual Growth Rate (CAGR) of 7.9% and 7.5% at the level of EU27 and EU28 respectively. More interestingly, the employment share and the intensity share components of the data professionals' indicator are also expected to significantly improve in 2017 and 2020 if compared to our estimates in 2016 (now estimated at 3% and 3.4% in 2007 and 2020 in the EU27 and 3.2% and 3.6% for the same years in the EU28), thus confirming the positive evolution of the workforce involved in data-related professions over the period under consideration.

An overview of Indicator 1 – Data Professionals is outlined in Table 15 below.

*Table 15: Indicator 1 — Data Professionals 2016-2017-2020 Baseline and Growth Rates*

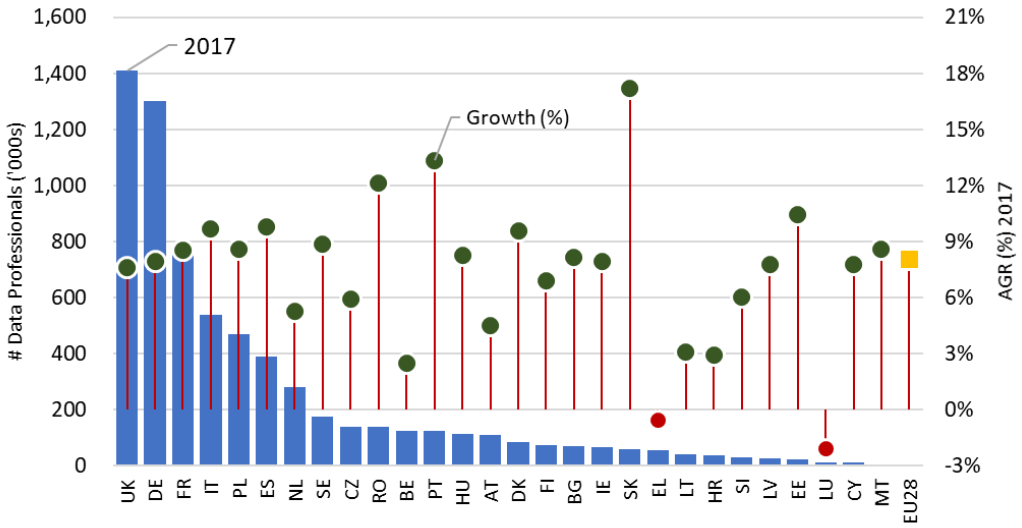
| Indicator 1 — Data Professionals 2016-2017-2020 Baseline and Growth Rates | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| N. | Region | Name | Description | 2016 | 2017 | 2020 Baseline | Growth 2017/2016 | CAGR 2020/2017 |
| 1.1 | EU27 | Number of data professionals | Total number of data professionals in EU (000s) | **4,875** | **5,273** | **6,619** | **8.2%** | **7.9%** |
| 1.1 | EU28 | Number of data professionals | Total number of data professionals in EU (000s) | **6,187** | **6,685** | **8,309** | **8.0%** | **7.5%** |
| 1.2 | EU27 | Employment share of data professionals | Share of data professionals on total employment in EU (%) | **2.8%** | **3.0%** | **3.4%** | 5.3% | 4.2% |
| 1.2 | EU28 | Employment share of data professionals | Share of data professionals on total employment in EU (%) | **3.1%** | **3.2%** | **3.6%** | 5.3% | 4.0% |
| 1.3 | EU27 | Intensity share of data professionals | Average number of data professionals per user company (units) | **9.6** | **10.2** | **12.3** | 5.9% | 6.3% |
| 1.3 | EU28 | Intensity share of data professionals | Average number of data professionals per user company (units) | **9.2** | **9.7** | **11.5** | 5.8% | 5.9% |

Source: European Data Market Monitoring Tool, IDC 2018

## 4.3 Data Professionals by Member State

The number of data professionals and the year-on-year growth rate in 2017 over 2016 and in 2020 according to the previous Baseline scenario by Member State are presented in Figure 6 and table 16 below. Not surprisingly, the absolute number of data professionals continues to be closely tied to the overall number of employment base and the size of the economy in each Member State, with the "Big 6" (U.K, Germany, France, Italy, Spain and Poland) continuing to feature the highest amount of data professionals in the EU in 2016, 2017 and 2020 according to the previous Baseline scenario.

*Figure 6: Data Professionals by Member State, 2017 (Units, '000, 2016-2017 Growth Rates %)*




Source: European Data Market Monitoring Tool, IDC 2018

The number of data professionals is consistent with the size and strength of the economies and industries in the Member States. For example, the U.K. has a strong data-based industry with Financial services and Professional services as key industries in the country. As a result, it has the largest number of data professionals among the Member States – as shown by the vertical bars. The major Member States dominate this chart. With a strong technology base seen supporting professionals in Germany and France. In Italy the strong retail presence contributes to the high number of data professionals, as the Retail industry is embracing digital transformation to maintain and grow its business. This contributes to Italy's above average growth, as shown by the green bubbles. Average growth is shown for reference by the yellow square at the right-hand end of the chart. The axis for the annual growth rate for data professionals is on the right-hand side of the chart.

*Table 16: Data Professionals by Member State, 2016-2017-2020 Baseline (Units, '000) and Growth Rates (%)*

| Member State | 2016 | 2017 | 2020 Baseline | Growth 2017/ 2016 | CAGR 2017/ 2020 |
|---|---|---|---|---|---|
| Austria | 104 | 108 | 116 | 4.5% | 2.2% |
| Belgium | 123 | 126 | 120 | 2.5% | -1.5% |
| Bulgaria | 66 | 71 | 84 | 8.1% | 5.5% |
| Croatia | 34 | 35 | 36 | 2.9% | 0,6% |
| Cyprus | 10 | 10 | 15 | 7.8% | 12.8% |
| Czech Republic | 132 | 139 | 160 | 5.9% | 4.8% |
| Denmark | 75 | 83 | 113 | 9.6% | 11.1% |
| Estonia | 19 | 21 | 30 | 10.5% | 11.7% |
| Finland | 70 | 74 | 90 | 6.9% | 6.7% |
| France | 725 | 787 | 1,051 | 8.5% | 10.1% |
| Germany | 1,204 | 1,300 | 1,622 | 7.9% | 7.7% |
| Greece | 56 | 56 | 65 | -0.6% | 5.2% |
| Hungary | 106 | 114 | 139 | 8.3% | 6.8% |
| Ireland | 60 | 65 | 81 | 8.0% | 7.4% |
| Italy | 490 | 538 | 701 | 9.7% | 9.3% |
| Latvia | 24 | 26 | 32 | 7.8% | 6.5% |
| Lithuania | 38 | 39 | 42 | 3.1% | 2.4% |
| Luxembourg | 12 | 12 | 10 | -2.1% | -6.4% |
| Malta | 5 | 6 | 8 | 8.6% | 10.8% |
| Netherlands | 266 | 280 | 307 | 5.3% | 3.1% |
| Poland | 431 | 468 | 573 | 8.6% | 7.0% |
| Portugal | 108 | 123 | 168 | 13.3% | 11.0% |
| Romania | 124 | 139 | 191 | 12.1% | 11.3% |
| Slovakia | 51 | 60 | 98 | 17.2% | 18.0% |
| Slovenia | 26 | 28 | 34 | 6.0% | 6.1% |
| Spain | 355 | 389 | 514 | 9.8% | 9.7% |
| Sweden | 161 | 175 | 218 | 8.9% | 7.7% |
| United Kingdom | 1,312 | 1,412 | 1,690 | 7.6% | 6.2% |
| **EU27** | **4,875** | **5,273** | **6,619** | **8.2%** | **7.9%** |
| **EU28** | **6,187** | **6,685** | **8,309** | **8.0%** | **7.5%** |

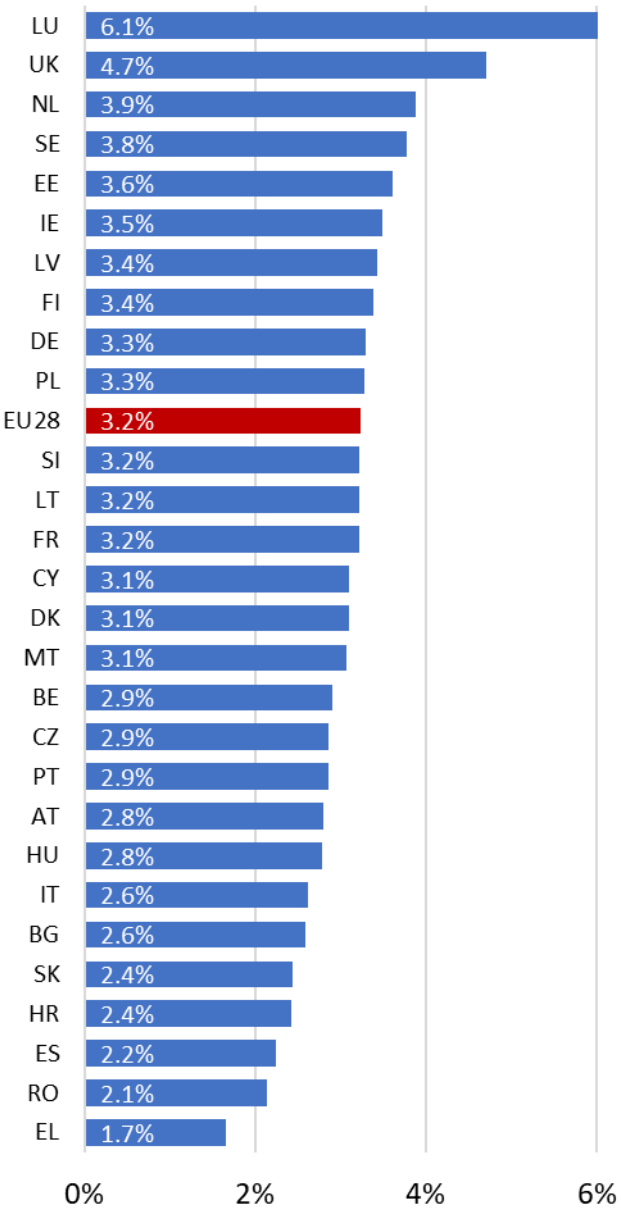
Source: European Data Market Monitoring Tool, IDC 2018

### 4.3.1 Employment Share of Data Professionals by Member State

Figure 7 below offers a detailed view by Member State of the data professionals' share on total employment in 2017 according to our most recent estimates. Both the EU27 and the EU28 average share of data professionals on total employment are estimated on a slight increase in 2017. In line with the results of the previous study, the countries with a dynamic data-based economy and a

significant share of ICT spending and innovation present a share of data professionals well above the EU average. This is the case of many small-to-medium-sized but data-driven Member States such as Luxembourg, the Netherlands, Sweden, Estonia, Ireland and Finland and the U.K. Indeed, the average share of data professionals on total employment in the EU is reduced of 0.2 percentage points in 2017 when excluding the United Kingdom.

*Figure 7:Share of Data Professionals on Total Employment by Member States, 2017 (%)*



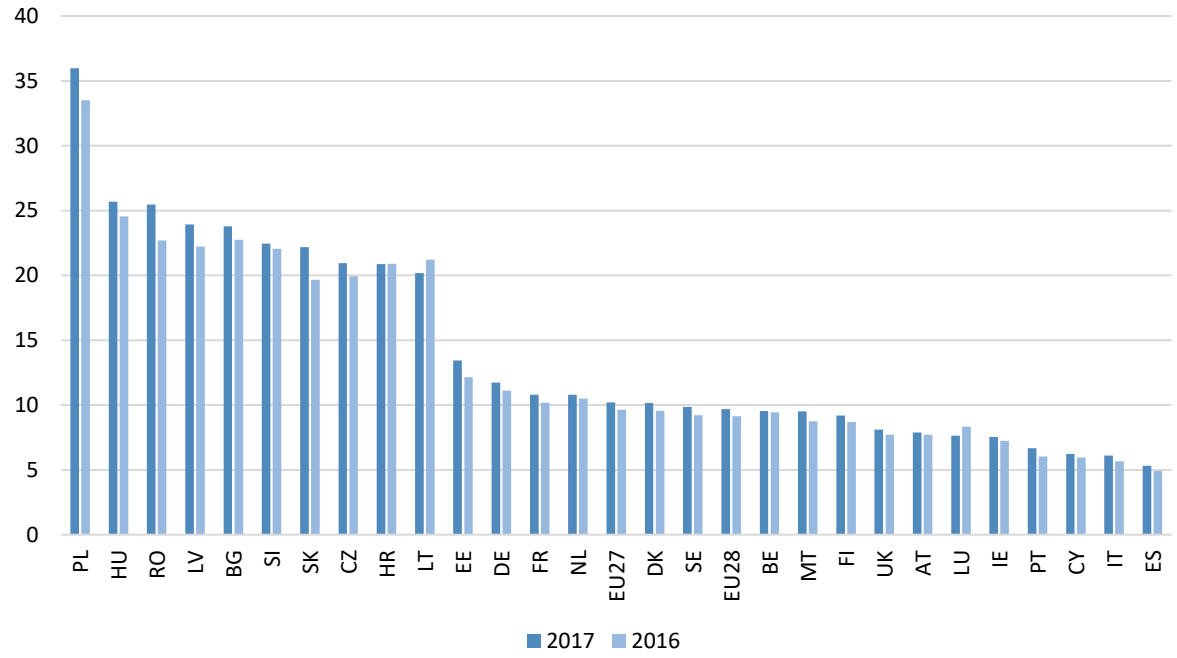

*Source: European Data Market Monitoring Tool, IDC 2018*

### 4.3.2 Intensity Share of Data Professionals by Member State

In 2017 the average number of data professionals by data user company has increased to 9.7 units in the EU28 and to 10.2 units in the EU27, a moderate improvement in line with the results obtained in the previous years - the same indicator was at 9.1 in 2014, 9.2 in 2015 and 2016 in the EU28 and 9.5 in 2014, 9.8 in 2015 and 9.6 in 2016 in the EU27. In line with the results obtained in the previous study, Member States from Central and Eastern Europe tend to exhibit a high or very high intensity

of data professionals. This is the case of Poland, Hungary, Romania, Bulgaria, Latvia and Lithuania, for example, where the overall number of data user companies is still lower than the EU average and the level of ICT spending is still limited if compared to Western European Member States. The consequence of this relative lack of data users and of technology investments tend to inflate the average number of data professionals by data user company.

*Figure 8: Average Number of Data Professionals by User Company, 2016-2017-2020 Baseline (Units)*




Source: European Data Market Monitoring Tool, IDC 2018

Data Sources: IDC estimates on Eurostat Labour Force Survey by Occupation and NACE II Industry Code 2018

## 4.4 Data Professionals by Industry

In absolute terms, data professionals continue to be spread out across all industries in the EU with a clear predominance in the Professional services, Retail & Wholesale, Manufacturing and the ICT sectors – all industries where the level of ICT investments and the size of the Data Market is already well developed (see Table 17 below). In terms of year-on-year growth, the most promising developments will come from traditional innovative sectors such as the ICT and the Professional services industries but also from the Education sector – all exhibiting above-average growth rates in 2017 over 2016 (Figure 9).

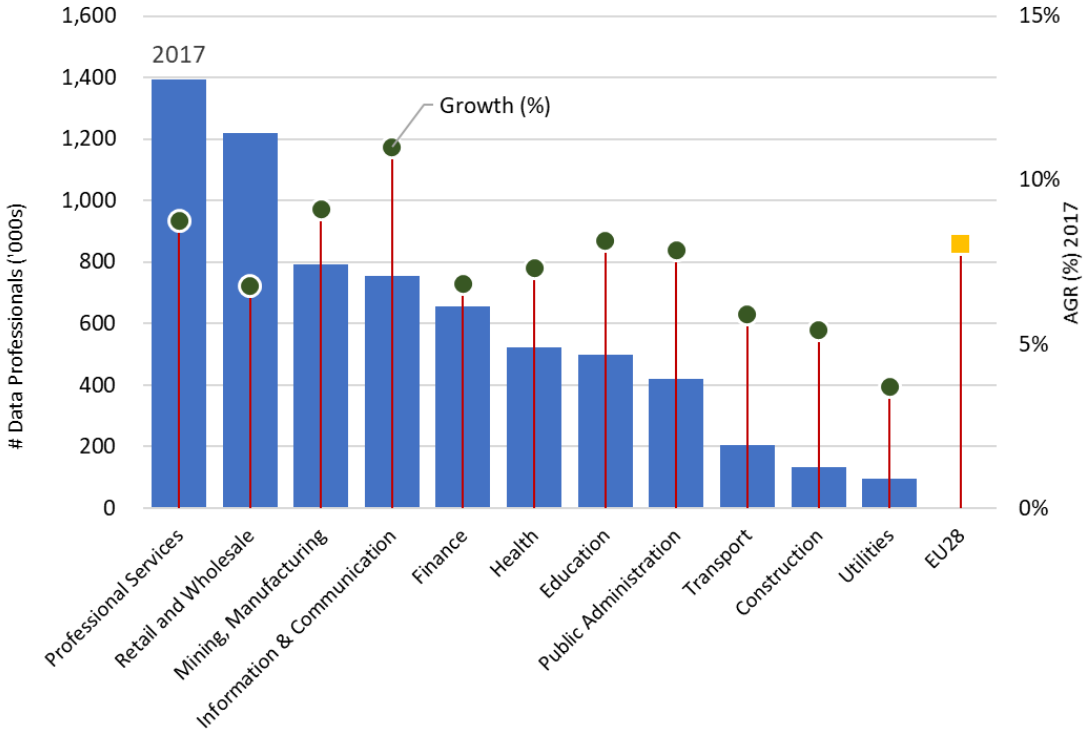
*Table 17: Data Professionals by Industry, 2016-2017-2020 Baseline (Units, '000) and Growth Rates (%)*

| Industry | 2016 | 2017 | 2020 Baseline | Growth 2017/2016 | CAGR 2017/2020 |
|---|---|---|---|---|---|
| Construction | 126 | 133 | 154 | 5.4% | 5.1% |
| Education | 459 | 497 | 595 | 8.2% | 6.2% |
| Finance | 615 | 657 | 818 | 6.8% | 7.6% |
| Health | 488 | 523 | 619 | 7.3% | 5.8% |
| Information & Communication | 680 | 755 | 1,017 | 11.0% | 10.4% |
| Mining, Manufacturing | 725 | 791 | 1,010 | 9.1% | 8.5% |
| Professional | 1,281 | 1,393 | 1,739 | 8.8% | 7.7% |

| Industry | 2016 | 2017 | 2020 Baseline | Growth 2017/2016 | CAGR 2017/2020 |
|---|---|---|---|---|---|
| services | | | | | |
| Public Administration | 388 | 418 | 521 | 7.9% | 7.6% |
| Retail and Wholesale | 1,141 | 1,219 | 1,485 | 6.8% | 6.8% |
| Transport | 192 | 203 | 237 | 5.9% | 5.3% |
| Utilities | 93 | 96 | 113 | 3.7% | 5.5% |
| EU28 | 6,187 | 6,685 | 8,309 | 8.0% | 7.5% |

Source: European Data Market Monitoring Tool, IDC 2018

*Figure 9: Data Professionals by Industry, 2017 (Units, '000) and Growth Rates (%)*




Source: European Data Market Monitoring Tool, IDC 2018

### 4.4.1 Employment Share of Data Professionals by Industry

Industries with a clear predominance of data-driven products and services are those that continue to show the highest levels of data professionals over the total numbers of employees by industry (Table 18). Not surprisingly, this continues to be the case of ICT, Finance and Professional services sectors - all sectors with an above-average penetration rate of data professionals in the period under consideration.

*Table 18: Share of Data Professionals on Total Employment by Industry, 2016-2017-2020 Baseline (%)*

| Industry | 2016 | 2017 | 2020 Baseline |
|---|---|---|---|
| Construction | 0.8% | 0.9% | 0.9% |
| Education | 2.7% | 2.9% | 3.2% |
| Finance | 9.4% | 9.9% | 11.6% |
| Health | 2.0% | 2.1% | 2.3% |
| Information & Communication | 10.1% | 10.8% | 12.7% |
| Mining, Manufacturing | 2.1% | 2.2% | 2.5% |

| Industry | 2016 | 2017 | 2020 Baseline |
|---|---|---|---|
| Professional services | 5.4% | 5.7% | 6.5% |
| Public Administration | 2.5% | 2.7% | 3.0% |
| Retail and Wholesale | 2.7% | 2.8% | 3.1% |
| Transport | 1.6% | 1.7% | 1.7% |
| Utilities | 2.8% | 3.0% | 3.5% |
| EU28 | 3.1% | 3.2% | 3.6% |

Source: European Data Market Monitoring Tool, IDC 2018

## 4.5 Data Professionals Forecasts at 2025

The number of data professionals in both the EU27 and EU28 is forecast to grow significantly under all the new three scenarios at 2025 as the use of data-driven innovation is expected to grow under unabatedly even under the less economically favourable scenario. A steady progression of the number of data professionals continues to emerge from our latest estimates. Under the new Baseline scenario, data professionals are expected to amount to 9.4 million in the EU27 and 11.5 million in the EU28 at 2025, thus representing a solid growth rate between 6.7% and 7.2% over the 2017-2025 period. In the new Challenge and High Growth scenarios, data professionals would be more than 8.6 million in the EU27 and 10.5 million in the EU28 and 10.9 million and 13,4 million respectively. Under all new scenarios, the CAGR over the period 2017-2025 would be in line with the CAGR featured by the Data Market growth, thus confirming again the close relationship between the two variables.

*Table 19: Data Professionals in 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) | 8,718 | 9,377 | 10,916 | 5.7% | 7.2% | 10.5% |
| 1.1 | EU28 | Number of data professionals | Total number of data professionals in EU (000s) | 10,564 | 11,489 | 13,450 | 4.9% | 6.7% | 10.1% |

Source: European Data Market Monitoring Tool, IDC 2018

### 4.5.1 Data Professionals Forecasts at 2025 – by Member State

At Member State level no significant variations are emerging in 2025 with respect to our estimates at 2020 under all three scenarios. The number of data professionals by Member State remains strongly correlated to the national size of the total employment, as well as to the number of data suppliers and data users by Member State. As a result, in 2025 the Member States with the highest number of data professionals will continue to be the "Big 6" (Germany, France, Italy, Poland, Spain, and the U.K.) followed by the Netherlands, Romania and Sweden in all the three scenarios under consideration (see Table 20 below).

*Table 20: Number of Data Professionals in 2025, Challenge, Baseline and High Growth Scenarios (Units, '000)*

| Country | 2025 Challenge | 2025 Baseline | 2025 High Growth | CAGR 2025/2020 Challenge Scenario (%) | CAGR 2025/2020 Baseline Scenario (%) | CAGR 2025/2020 High Growth Scenario (%) |
|---|---|---|---|---|---|---|
| Austria | 128 | 127 | 160 | 2.0% | 2.0% | 6.7% |
| Belgium | 122 | 127 | 151 | 0.3% | 1.1% | 4.6% |
| Bulgaria | 102 | 113 | 120 | 4.1% | 6.1% | 7.5% |
| Croatia | 43 | 49 | 52 | 3.5% | 6.3% | 7.4% |
| Cyprus | 23 | 25 | 28 | 9.5% | 11.1% | 14.0% |
| Czech Republic | 217 | 216 | 251 | 6.3% | 6.2% | 9.4% |
| Denmark | 166 | 176 | 217 | 8.0% | 9.1% | 13.9% |
| Estonia | 41 | 46 | 56 | 6.5% | 9.1% | 13.2% |
| Finland | 127 | 127 | 134 | 7.1% | 7.1% | 8.2% |
| France | 1,523 | 1,557 | 1,937 | 7.7% | 8.2% | 13.0% |
| Germany | 2,307 | 2,379 | 2,662 | 7.3% | 8.0% | 10.4% |
| Greece | 95 | 97 | 115 | 7.9% | 8.3% | 12.1% |
| Hungary | 160 | 184 | 185 | 2.8% | 5.8% | 5.8% |
| Ireland | 105 | 107 | 116 | 5.4% | 5.8% | 7.4% |
| Italy | 845 | 993 | 1,277 | 3.8% | 7.2% | 12.7% |
| Latvia | 45 | 46 | 57 | 7.2% | 7.9% | 12.2% |
| Lithuania | 55 | 54 | 69 | 5.3% | 5.0% | 10.5% |
| Luxembourg | 13 | 14 | 16 | 6.1% | 7.5% | 10.7% |
| Malta | 11 | 12 | 15 | 8.0% | 9.8% | 13.5% |
| Netherlands | 314 | 394 | 406 | 0.5% | 5.1% | 5.7% |
| Poland | 684 | 748 | 863 | 3.6% | 5.4% | 8.5% |
| Portugal | 218 | 244 | 266 | 5.3% | 7.7% | 9.6% |
| Romania | 284 | 290 | 300 | 8.2% | 8.7% | 9.4% |
| Slovakia | 134 | 170 | 200 | 6.4% | 11.6% | 15.2% |
| Slovenia | 39 | 48 | 51 | 3.3% | 7.6% | 8.7% |
| Spain | 692 | 748 | 947 | 6.1% | 7.8% | 13.0% |
| Sweden | 221 | 284 | 269 | 0.3% | 5.4% | 4.3% |
| United Kingdom | 1,846 | 2,111 | 2,534 | 1.8% | 4.5% | 8.4% |
| **EU27** | **8,718** | **9,377** | **10,916** | **4.9%** | **6.7%** | **10.1%** |
| **EU28** | **10,564** | **11,489** | **13,450** | 2.0% | 2.0% | 6.7% |

Source: European Data Market Monitoring Tool, IDC 2018

Table 20 shows the forecast for number of data professionals under each of the Challenge, Baseline, and High Growth scenarios. The United Kingdom is expected to have left the European Union by 2025, but the country makes a significant contribution to overall growth in the number of data professionals in the EU because of its already strong presence in data oriented industries. These include Financial services and Professional services – both of which are early adopters of data and

data tools to manage and grow their business. In both the Challenge and High Growth scenarios the U.K. lists the growth rate notably.

Some of the Member States show what might be considered a remarkable growth when compared to data intense countries such as the U.K. or Germany. But this is primarily due to the low level of data professionals in the Member States, and the relatively small population, from which is it easy to show what might be considered unusually high growth. Caution should be exercised with these high-growth-but-low-population countries.

The Nordic countries have shown in the past to be early adopters of technology and this is apparent in the growth and size data in the scenarios. These countries will build on their well-educated workforce to benefit from digital transformations, and the higher concentration of digital devices compared with other Member States. Finally, Estonia deserves a mention as a special case, as it embraced the whole ethos of digital technology from an early stage, and continues to support the introduction of digital technology in all areas of government, daily life, and business.

## 4.6 Key Findings

The number of workers who collect, store, manage, analyse and visualise data as their primary or as a relevant part of their activity is estimated at 5.3 million in the EU27 and 6.7 million in the EU28 in 2017, thus marking a significant increase in 2017 over the previous year (8.2% and 8% year-on-year respectively).

The employment share and the intensity share components of the data professionals' indicator are also expected to significantly improve in 2017 and 2020 if compared to our estimates in 2016 (now estimated at 3% and 3.4% in 2017 and 2020 in the EU27 and 3.2% and 3.6% for the same years in the EU28), thus confirming the positive evolution of the workforce involved in data-related professions over the period under consideration. This result shows direct continuity with the findings obtained in the previous European Data Market Study, SMART 2013/0063.

The average number of data professionals per company will continue to grow steadily (although slowly) in 2017 and is estimated at an average of 11.5 per company by 2020 in the EU28.

The number of data professionals in the EU28 is expected to rise by a compound rate of 6.7% by 2025. This growth could be as high as 10.1% per year if the highest growth scenario is achieved – giving nearly 13.5 million data professionals in the EU28 and almost 11 million data professionals in the EU27 at 2025.

There is an even balance of data professionals' penetration across all Member States in 2017, aside from some unsurprising outliers which are data-intense economies (such as the U.K., Luxembourg or the Netherlands), or economies still struggling to make the most out of the ongoing digitisation process (e.g.: Romania, Greece).

The highest concentration of data professionals is in the Professional services industry, followed by Retail and Wholesale. There is a notably wider spread of the number of data professionals by industry, although there is a fair correlation between the number of companies and employees in those industries and the number of data user companies in the same industries.

The highest growth in data professionals' year-on-year (in 2017 over 2016) is in the Information and Communications industry – although the growth is fairly even across all industries as each can benefit from transforming its business into a digital business.

# 5. MEASURING THE DATA COMPANIES

## 5.1 Definition

***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.*

Indicator 2 measures separately:

- European data suppliers, counted as legal entities based in one EU Member State, as a share of the total number of enterprises included in the Information and Communication industry and Professional services industry classification (J and M in the NACE rev2).
- European data users, counted as legal entities based in one EU Member State, as a share of the total number of private enterprises in the EU.

*Table 21: Data Sources: Data Companies (Data Suppliers and Data Users)*

| Data Source | Updated |
|---|---|
| Eurostat annual structural business statistics | Dec-2017 |
| Eurostat Business Demographic Statistics | Dec-2017 |
| IDC Worldwide Black Book (standard edition) | Nov-2017 |

## 5.2 Measuring Data Companies

Table 22 provides an overview of our latest estimates of the number and share of data supplier and data user companies in the EU27 and EU28 in the years 2016, 2017 and 2020 Baseline scenario as defined in the European Data Market Study (SMART 2013/0063).

The numbers of data suppliers continue to grow at a faster pace than the numbers of data users: the former are estimated at almost 142,000 in the EU27 and at more than 276,000 units in the EU28, thus exhibiting a year-on-year growth of 5.7% in 2017. Data users, instead, are projected to grow at 2.1% in 2017, amounting to almost 516,000 in the EU27 and to 690,650 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 more dynamic pictures of data companies in the EU, with growth rates constantly increasing over the past four years.

This positive dynamic is reflected in the percentage shares 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 now estimated at 11.5% in the EU27 (up 0.6 percentage points vis-à-vis the previous year) and 15% in the EU28 (up 0.8 percentage points), while the data users' penetration rates are up a modest 0.1 percentage point in 2017 in both the EU27 and EU28.

*Table 22: Indicator 2 – Data Companies, 2016-2017-2020 Baseline and Growth Rates*

| N. | Name | Description | Market | 2016 | 2017 | 2020 | Growth 2017/2016 | CAGR 2020/2017 |
|---|---|---|---|---|---|---|---|---|
| **2.1** | Number of data suppliers | Total number of data suppliers measured as legal entities based in the EU (000s) | EU27 | 134,300 | 141,900 | 157,150 | 5.7% | 3.5% |
| **2.1** | Number of data suppliers | Total number of data suppliers measured as legal entities based in the EU (000s) | EU28 | 261,450 | 276,450 | 305,600 | 5.7% | 3.4% |
| **2.2** | Share of data suppliers | % share of data companies on total companies in the ICT and Professional services industries | EU27 | 10.9% | 11.5% | 12.8% | 5.6% | 3.5% |
| **2.2** | Share of data suppliers | % share of data companies on total companies in the ICT and Professional services industries | EU28 | 14.2% | 15.0% | 16.5% | 5.5% | 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 | 516,550 | 539,800 | 2.1% | 1.5% |
| **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 | 690,650 | 721,850 | 2.1% | 1.5% |
| **2.4** | Share of data users | % share of data users on total companies in the EU industry | EU27 | 5.7% | 5.8% | 6.1% | 1.8% | 1.6% |
| **2.4** | Share of data users | % share of data users on total companies in the EU industry | EU28 | 6.5% | 6.6% | 6.9% | 1.9% | 1.6% |

Source: European Data Market Monitoring Tool, IDC 2018

## 5.3 Measuring Data Suppliers

### 5.3.1 Data Suppliers by Member State

The number of data suppliers (as shown in Table 22 above) continues to grow, but not at the same rate as the Data Market as a whole (See chapter 7). The European Market will grow at a compound rate of 3.9% between 2016 and 2020, with the highest growth seen in Cyprus and the lowest in Croatia. However, the Data Market and population of industries in both these countries is very small compared with five major economies', and among these, growth is closer to the average for the

European Union as a whole. The number of data companies reflects the inherent business focus, and the strong financial, and services industries in the UK make the number of data suppliers in this member state nearly five times that of the next biggest supplier of data tools and services.

*Table 23: Data Suppliers by Member State, 2016-2017-2020 Baseline (Units, '000)*

| Member State | 2016 | 2017 | 2020 |
|---|---|---|---|
| Austria | 4,250 | 4,550 | 4,750 |
| Belgium | 2,350 | 2,450 | 2,950 |
| Bulgaria | 1,150 | 1,200 | 1,350 |
| Croatia | 700 | 700 | 800 |
| Cyprus | 400 | 500 | 550 |
| Czech Republic | 1,950 | 2,050 | 2,200 |
| Denmark | 3,650 | 3,900 | 4,250 |
| Estonia | 450 | 500 | 550 |
| Finland | 2,850 | 2,950 | 3,100 |
| France | 12,300 | 12,950 | 13,900 |
| Germany | 25,500 | 27,100 | 29,250 |
| Greece | 5,450 | 5,700 | 6,200 |
| Hungary | 3,250 | 3,450 | 3,850 |
| Ireland | 3,350 | 3,600 | 3,800 |
| Italy | 18,450 | 19,450 | 21,800 |
| Latvia | 550 | 550 | 650 |
| Lithuania | 600 | 650 | 750 |
| Luxembourg | 450 | 500 | 550 |
| Malta | 250 | 250 | 250 |
| Netherlands | 5,300 | 5,650 | 6,800 |
| Poland | 5,650 | 5,900 | 6,500 |
| Portugal | 3,950 | 4,150 | 4,550 |
| Romania | 5,450 | 5,750 | 5,850 |
| Slovakia | 2,150 | 2,300 | 2,700 |
| Slovenia | 500 | 550 | 650 |
| Spain | 15,250 | 16,000 | 17,350 |
| Sweden | 8,150 | 8,600 | 11,250 |
| United Kingdom | 127,150 | 134,550 | 148,450 |
| EU27 | 134,300 | 141,900 | 157,150 |
| EU28 | 261,450 | 276,450 | 305,600 |

Source: European Data Market Monitoring Tool, IDC 2018

### 5.3.2 Data Suppliers by Industry

As in the previous European Data Market Study (SMART 2013/0063), we have identified data suppliers as belonging to two specific industry sectors: Information and Communication technologies and Professional services. The distribution of data suppliers by these two industries remains essentially stable with the ICT sector growing slightly faster than the Professional services sector in 2017 over the previous year.

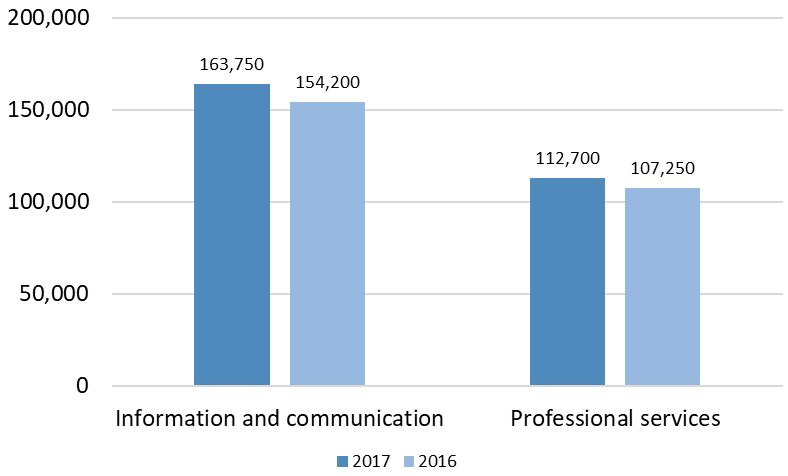
*Table 24: Data Suppliers by Industry, 2016-2017-2020 Baseline (Units, '000)*

| Industry | 2016 | 2017 | 2020 Baseline | Growth 2017/2016 |
|---|---|---|---|---|
| Information and Communication | 154,200 | 163,750 | 180,200 | 6.2% |
| Professional services | 107,250 | 112,700 | 125,400 | 5.1% |
| EU28 | 261,450 | 276,450 | 305,600 | 5.7% |

Source: European Data Market Monitoring Tool, IDC 2018

Figure 10 shows the relative sizes of the ICT and Professional services industries as suppliers into the Data Market, and both appear strong, growing at similar rates in 2017.

*Figure 10: Data Suppliers by Industry, EU28, 2016-2017 (Units)*




Source: European Data Market Monitoring Tool, IDC 2018

### 5.3.3 Data Suppliers by Company Size

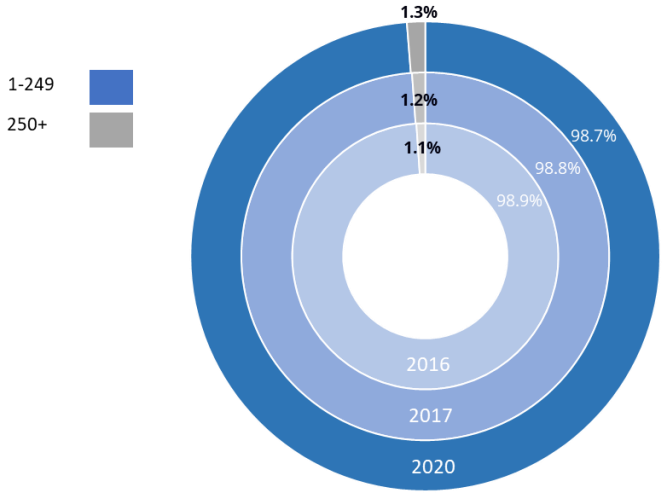
The number of companies by size is completely dominated by those in the 1-249 size band, but this reflects the general mix of business size in Europe, with over 99% of companies in the 1-249 size band.

*Table 25: Data Suppliers by Company Size, 2016-2017-2020 Baseline (Units, '000)*

| Size Band | 2016 | 2017 | 2020 Baseline |
|---|---|---|---|
| 1-249 | 258,650 | 273,250 | 301,775 |
| 250+ | 2,800 | 3,200 | 3,825 |
| Grand Total | 261,450 | 276,450 | 305,600 |

Source: European Data Market Monitoring Tool, IDC 2018

*Figure 11: Data Suppliers by Company Size, EU28, 2016-2017- 2020 Baseline (%)*



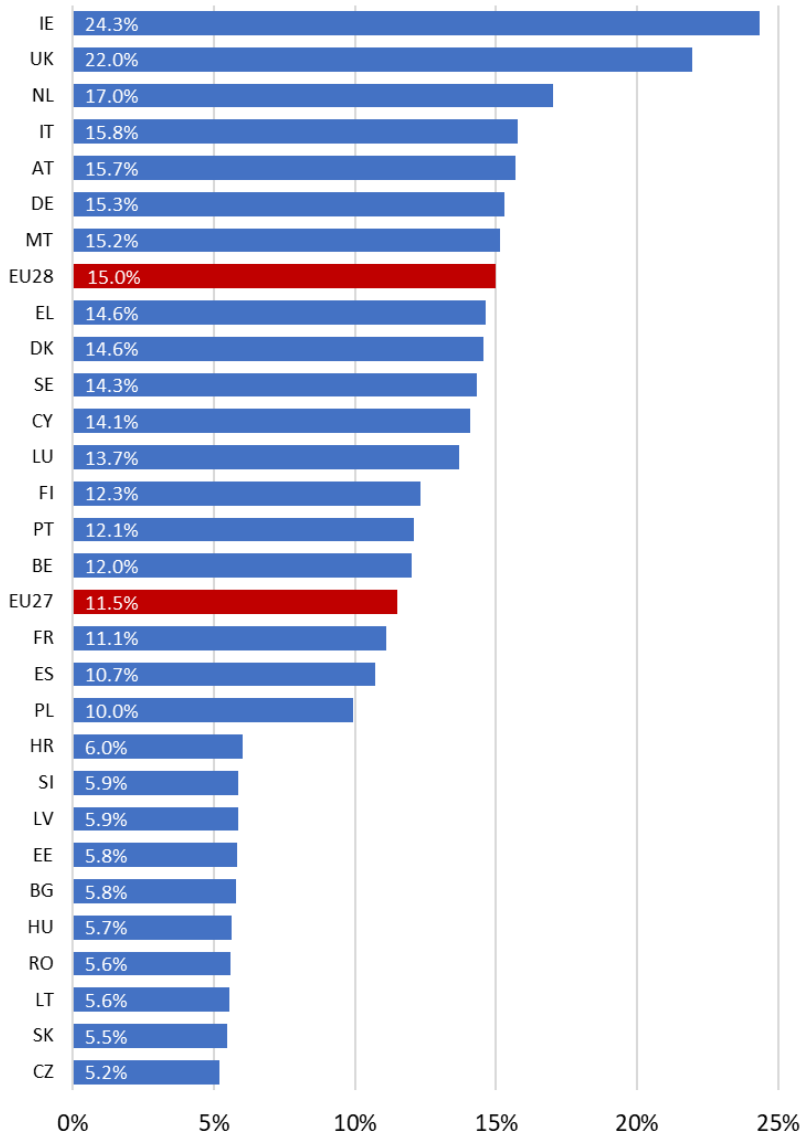

Source: European Data Market Monitoring Tool, IDC 2018

### 5.3.4 Share of Data Suppliers on Total Companies

The share of data suppliers on the total of data companies (i.e. the percentage share of data supplier companies on the total companies in the ICT and Professional services industries) continues to grow.

It was 13.6% in the EU28 in 2014 and it is now estimated at 15% (10.7% in 2014 and 11.5% in 2017 in the EU27). With the UK clearly dominating the data suppliers' scene in Europe, the EU27 would exhibit a much lower penetration rate of data suppliers than the EU28.

*Figure 12: Share of Data Suppliers on Total J and M Sectors, 2017 (%)*




Source: European Data Market Monitoring Tool, IDC 2018

### 5.3.5 Data Suppliers Forecasts at 2025

The outlook for data suppliers is continued growth, but the baseline growth to 2025 does not fully match the growth seen to 2020 according to our previous Baseline scenario – primarily because the market is larger and so growth becomes more difficult. In spite of 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 individuals 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.

Nonetheless, innovation often comes from smaller organizations, whose focus is on this emerging and high valued market – but again these smaller companies tend to be acquired by the larger ones as their technology becomes understood – and for many this is the exit strategy for the company founders.

*Table 26: Data Suppliers Forecast 2025 by Member State - Three Scenarios (Units; '000); CAGR 2025-2020 (%)*

| Member State | 2025 Challenge | 2025 Baseline | 2025 High Growth | CAGR 2025/2020 Challenge Scenario (%) | CAGR 2025/2020 Baseline Scenario (%) | CAGR 2025/2020 High Growth Scenario (%) |
|---|---|---|---|---|---|---|
| Austria | 5,100 | 5,300 | 6,100 | 1.4% | 2.2% | 5.1% |
| Belgium | 3,300 | 3,550 | 4,150 | 2.3% | 3.8% | 7.1% |
| Bulgaria | 1,600 | 1,700 | 1,750 | 3.5% | 4.7% | 5.3% |
| Croatia | 900 | 950 | 1,000 | 2.4% | 3.5% | 4.6% |
| Cyprus | 650 | 650 | 650 | 3.4% | 3.4% | 3.4% |
| Czech Republic | 2,250 | 2,500 | 2,850 | 0.5% | 2.6% | 5.3% |
| Denmark | 4,650 | 4,900 | 5,600 | 1.8% | 2.9% | 5.7% |
| Estonia | 600 | 650 | 750 | 1.8% | 3.4% | 6.4% |
| Finland | 3,300 | 3,450 | 3,700 | 1.3% | 2.2% | 3.6% |
| France | 15,300 | 15,550 | 17,850 | 1.9% | 2.3% | 5.1% |
| Germany | 31,900 | 32,850 | 35,950 | 1.7% | 2.3% | 4.2% |
| Greece | 6,400 | 7,050 | 7,250 | 0.6% | 2.6% | 3.2% |
| Hungary | 4,500 | 4,750 | 5,050 | 3.2% | 4.3% | 5.6% |
| Ireland | 4,300 | 4,450 | 4,800 | 2.5% | 3.2% | 4.8% |
| Italy | 21,950 | 24,250 | 28,950 | 0.1% | 2.2% | 5.8% |
| Latvia | 650 | 750 | 750 | 0.0% | 2.9% | 2.9% |
| Lithuania | 800 | 900 | 900 | 1.3% | 3.7% | 3.7% |
| Luxembourg | 650 | 650 | 750 | 3.4% | 3.4% | 6.4% |
| Malta | 300 | 350 | 350 | 3.7% | 7.0% | 7.0% |
| Netherlands | 7,750 | 8,450 | 9,150 | 2.6% | 4.4% | 6.1% |
| Poland | 7,150 | 8,000 | 9,150 | 1.9% | 4.2% | 7.1% |
| Portugal | 4,800 | 5,150 | 5,450 | 1.1% | 2.5% | 3.7% |
| Romania | 6,300 | 6,550 | 7,500 | 1.5% | 2.3% | 5.1% |
| Slovakia | 3,100 | 3,500 | 4,000 | 2.8% | 5.3% | 8.2% |
| Slovenia | 700 | 750 | 750 | 1.5% | 2.9% | 2.9% |
| Spain | 17,800 | 19,650 | 22,300 | 0.5% | 2.5% | 5.1% |
| Sweden | 13,400 | 14,800 | 15,350 | 3.6% | 5.6% | 6.4% |
| United Kingdom | 168,600 | 175,900 | 208,100 | 2.6% | 3.5% | 7.0% |
| EU27 | 170,100 | 182,050 | 202,800 | 1.6% | 3.0% | 5.2% |
| EU28 | 338,700 | 357,950 | 410,900 | 2.1% | 3.2% | 6.1% |

Source: European Data Market Monitoring Tool, IDC 2018

Table 26 presents the three scenarios for the forecast for data suppliers, and as previously mentioned the expectation is for slightly lower growth as the market moves towards mainstream. The United Kingdom dominates this table with its strong use and hence supply of data technology in the key data industries. The restricting of data suppliers to Professional services and Information and Communication enterprises builds on the already strong infrastructure there. In addition, the U.K. has a large share of small and medium businesses in Professional services, which adds to its already large total of data supplier organisations. However, Germany also has a robust infrastructure in communications, which makes it the second largest Member State in terms of data suppliers. There

is not significant variation in growth across the Member States because the future of business will come from digital transformation – with some organisations predicting perhaps as much as 55 percent of business will be digital over the next few years – so all Member States will continue to support investment and growth in the digital economy. At this point the cost of entry is also low, which makes it easier for new entrants. However, as digital technology becomes more advanced we anticipate the ease of entry to this market will be reduced.

We anticipate the number of data suppliers will grow at a lower rate than the total digital market though, as each organisation will aim to grow its revenues and profits, and as a result the number of new data suppliers will not keep pace with the growth in the Data Market.

## 5.4 Measuring Data Users

### 5.4.1 Data Users by Member State

While there are significantly more data user companies than data suppliers, the growth in the number of these companies is slower – with the exception of the 250+ size band. Larger companies have the capacity and greater willingness to embrace Digital transformation because they understand the benefits to digitizing their business, and can dedicate teams to this. It is more likely that many smaller companies will never digitise, particularly in industries such as Construction because of the nature of their business, so growth in the smaller size bands is slower. Adoption rates as seen in user company surveys in 2016 and 2017 is slower among smaller companies because many of them find it difficult to understand what is involved and what are the benefits to their business, so these companies need more education and support if they are to change the way they conduct business and become more data oriented.

*Table 27: Data Users by Member State, 2016-2017-2020 Baseline (Units, '000); Growth rate 2017-2016 (%)*

| Member State | 2016 | 2017 | 2020 Baseline | Growth 2017/2016 |
|---|---|---|---|---|
| Austria | 13,450 | 13,750 | 14,100 | 2.2% |
| Belgium | 13,000 | 13,200 | 14,450 | 1.5% |
| Bulgaria | 2,900 | 3,000 | 3,150 | 3.4% |
| Croatia | 1,650 | 1,700 | 1,700 | 3.0% |
| Cyprus | 1,600 | 1,650 | 1,750 | 3.1% |
| Czech Republic | 6,600 | 6,650 | 6,750 | 0.8% |
| Denmark | 7,900 | 8,150 | 8,350 | 3.2% |
| Estonia | 1,600 | 1,600 | 1,800 | 0.0% |
| Finland | 8,000 | 8,100 | 8,200 | 1.3% |
| France | 71,300 | 72,850 | 74,700 | 2.2% |
| Germany | 108,300 | 110,900 | 114,450 | 2.4% |
| Greece | 10,700 | 10,900 | 11,150 | 1.9% |
| Hungary | 4,300 | 4,450 | 4,700 | 3.5% |
| Ireland | 8,350 | 8,650 | 8,700 | 3.6% |
| Italy | 86,600 | 88,150 | 92,950 | 1.8% |
| Latvia | 1,100 | 1,100 | 1,200 | 0.0% |
| Lithuania | 1,800 | 1,950 | 1,950 | 8.3% |
| Luxembourg | 1,450 | 1,550 | 1,600 | 6.9% |
| Malta | 600 | 600 | 700 | 0.0% |

| Member State | 2016 | 2017 | 2020 Baseline | Growth 2017/2016 |
|---|---|---|---|---|
| Netherlands | 25,350 | 26,000 | 28,800 | 2.6% |
| Poland | 12,850 | 13,000 | 13,600 | 1.2% |
| Portugal | 17,950 | 18,400 | 19,150 | 2.5% |
| Romania | 5,450 | 5,450 | 5,450 | 0.0% |
| Slovakia | 2,600 | 2,700 | 2,850 | 3.8% |
| Slovenia | 1,200 | 1,250 | 1,350 | 4.2% |
| Spain | 71,950 | 73,100 | 75,600 | 1.6% |
| Sweden | 17,400 | 17,750 | 20,650 | 2.0% |
| United Kingdom | 170,200 | 174,100 | 182,050 | 2.3% |
| EU27 | 505,950 | 516,550 | 539,800 | 2.1% |
| EU28 | 676,150 | 690,650 | 721,850 | 2.1% |

Source: European Data Market Monitoring Tool, IDC 2018

In Table 27 data users represent those companies which will use data as a source for their decision making. However, many organisations are not yet familiar with the availability of data sources, or do not have the expertise to exploit these data sources in their decision making. IDC[13] research shows adoption of big data tools among organisations is slowing, particularly among smaller companies, and a driving factor is having the expertise to implement these tools – they can be complex and difficult to manage.

Values in this table are rounded to the nearest 50 to reflect that this is forecast data (2017 published data is not yet available), and indicates the degree of accuracy in the data. However, for smaller countries this rounding can exaggerate growth, where the number of data users is small.

The U.K. again dominates this table, accounting for 25 percent of data user companies. This Member State's focus on data technology segments such as Finance and Professional services accounts for this, together with its share of total EU companies. Its focus means the number of data companies will grow above average by 2020.

The larger Member States (U.K., Germany, France, Italy) account for 50 percent of data user companies by 2020, with the remaining 24 accounting for the other 50 percent. Again, it is working population that gives these four countries their share, but as mentioned previously, digital transformation needs considerable expertise, and it is larger companies that can afford to invest now in the technology and change required to make this transformation. Future growth is likely to come from a combination of these larger companies embracing digital transformation fully, and emerging enterprises that implement digital technology to seize these new business opportunities.

### 5.4.2 Data Users by Industry

The distribution of data users by industry in 2016 and 2017 presents the same characteristics outlined in our previous study for the years 2013-2015 – Professional services dominates the scene, followed by Manufacturing, Retail & Wholesale, ICT, and Transport.

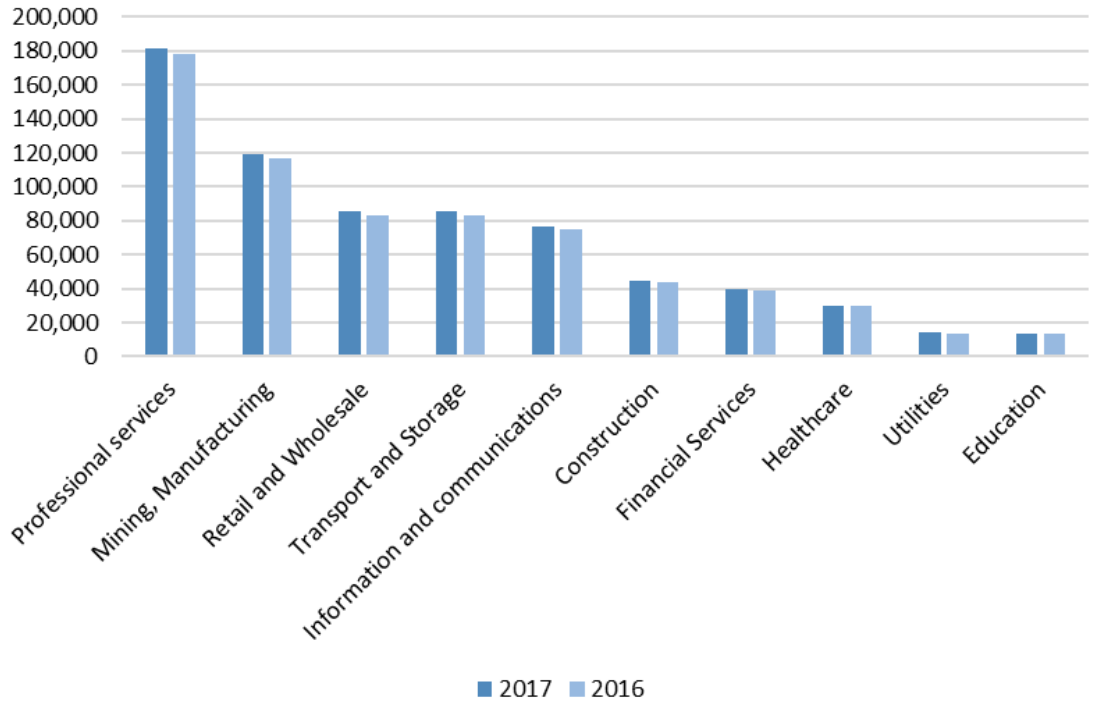
[13] IDC Market Forecast: "Worldwide Big Data and Analytics Software Forecast, 2017–2021, Market Forecast", DOC # US42891017, IDC July 2017

*Table 28: Data Users by Industry, 2016-2017-2020 Baseline (Units, '000); Growth Rate 2017-2016 (%)*

| Industry | 2016 | 2017 | 2020 Baseline | Growth 2017/2016 |
|---|---|---|---|---|
| Construction | 44,000 | 44,750 | 46,250 | 1.7% |
| Education | 13,550 | 13,800 | 14,250 | 1.8% |
| Financial services | 38,900 | 39,850 | 41,600 | 2.4% |
| Healthcare | 29,500 | 30,150 | 31,200 | 2.2% |
| Information and Communication | 74,550 | 76,550 | 80,200 | 2.7% |
| Mining, Manufacturing | 116,700 | 119,300 | 124,150 | 2.2% |
| Professional services | 178,550 | 181,750 | 192,250 | 1.8% |
| Retail and Wholesale | 83,200 | 85,350 | 88,750 | 2.6% |
| Transport and Storage | 83,450 | 85,150 | 88,500 | 2.6% |
| Utilities | 13,750 | 14,000 | 14,700 | 2.0% |
| **EU28** | **676,150** | **690,650** | **721,850** | **2.1%** |

Source: European Data Market Monitoring Tool, IDC 2018

*Figure 13: Data Users by Industry, 2016-2017 (Units)*




Source: European Data Market Monitoring Tool, IDC 2018

Table 28 and Figure 13 make clear the importance of the key industries – with Professional services being the industry with the largest number of data users. However, Manufacturing and Retail have already invested in digital transformation to gain efficiencies of digital technologies, and to grow business. IDC's annual end user survey for 2017 shows Retail and Wholesale is already setting its priorities towards business growth and marketing through the use of digital technology such as big data and analytics, while Manufacturing is focusing on improved efficiency through monitoring and tracking of equipment and materials.

### 5.4.3 Data Users by Company Size

In line with the overall structure of the EU industry, the SMEs' share of data users remains very high in Europe with small and medium data users' companies continuing to represent 98.8% of the overall data users' population in the EU28.

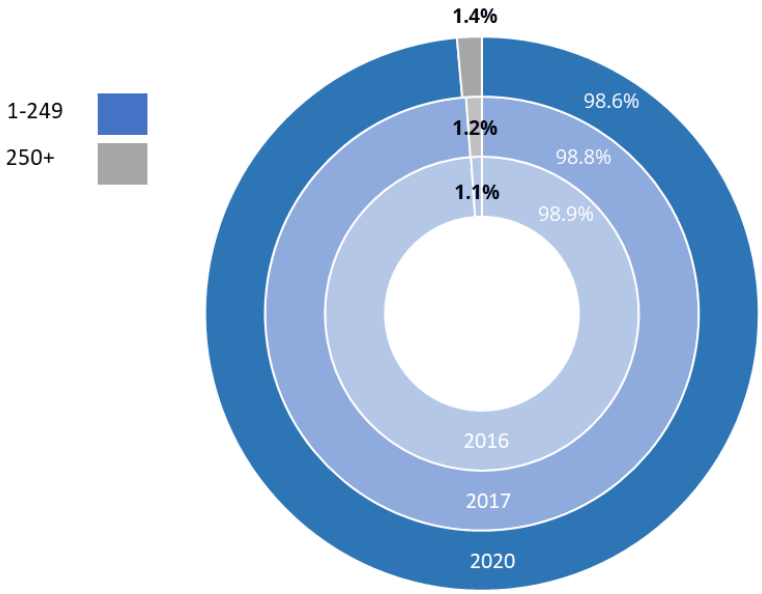
*Table 29: Data Users by Company Size, 2016-2017-2020 Baseline (Units; '000); Growth Rate 2017-2016 (%)*

| Size Band | 2016 | 2017 | 2020 Baseline | Growth 2017/2016 |
|---|---|---|---|---|
| 1-249 | 668,610 | 682,590 | 711,450 | 2.1% |
| 250+ | 7,540 | 8,060 | 10,400 | 6.9% |
| EU28 | 676,150 | 690,650 | 721,850 | 2.1% |

Source: European Data Market Monitoring Tool, IDC 2018

From table 29 it is clear that growth is greater among larger enterprises. This reflects IDC's user survey data which shows adoption of digital technologies is highest among larger companies. These companies can invest in the resources needed to understand and implement digital transformation. While smaller companies – particularly those with fewer than 100 employees struggle to make these investments. Digital technology is complex and needs considerable understanding to be well implemented. Indeed, smaller companies need constant support from technology suppliers to help understand this. In addition, many smaller companies are technology followers, not technology leaders. As a result, they tend to look for case studies where digital solutions were implemented well, and as this is early in the development cycle there are few cases to offer.

*Figure 14: Data User Companies by Company Size, 2016-2017-2020 Baseline (%)*




Source: European Data Market Monitoring Tool, IDC 2018

### 5.4.4 Data Users Average Spend by Company Size

While certainly less significant in numbers, medium and large companies continue to display a much higher average spending in data products and services with respect to SMEs. In 2017, data users of 250 employees and more will spend on average 5.6 million Euros versus only 30,000 Euros spent by small and medium companies.

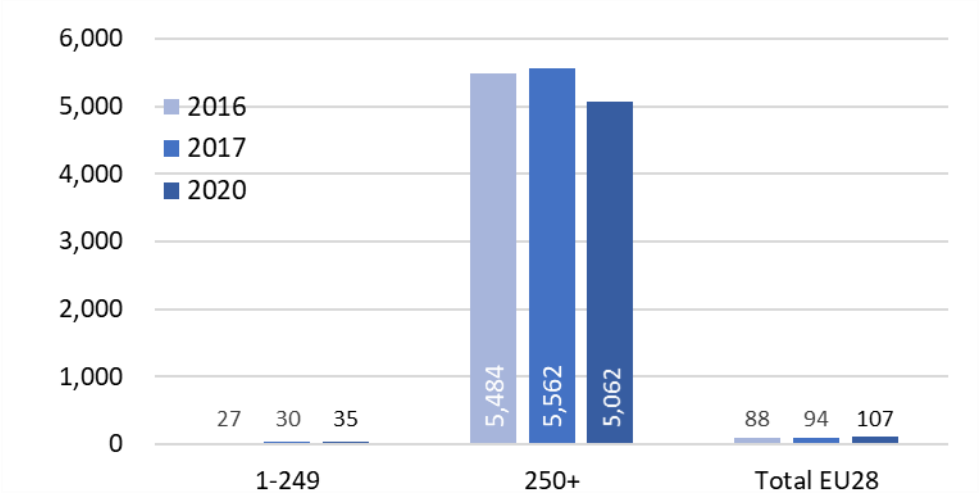
*Table 30: Data Users' Average Spend by Company Size, 2016-2017-2020 Baseline (Units; '000); Growth Rate 2017-2016 (%)*

| Size Band | 2016 | 2017 | 2020 Baseline | Growth 2017/2016 |
|---|---|---|---|---|
| 1-249 | 27 | 30 | 35 | 8.7% |
| 250+ | 5,484 | 5,562 | 5,062 | 1.4% |
| Total EU28 | 88 | 94 | 107 | 6.9% |

Source: European Data Market Monitoring Tool, IDC 2018

Table 30 reinforces the finding that the investment made by larger enterprises are way higher than those made by SMEs. Large organizations spend more than 100 times the investment of smaller companies. This shows the level of investment required, although this will reduce as competition increases, and the different technologies become cheaper. Larger organizations will of course spend more as they have a larger working population, but even so, smaller companies are unwilling or unable to make the investments needed to fully understand and implement digital technologies.

*Figure 15: Average Spend per Company by Company Size, 2016-2017-2020 Baseline (€, '000)*

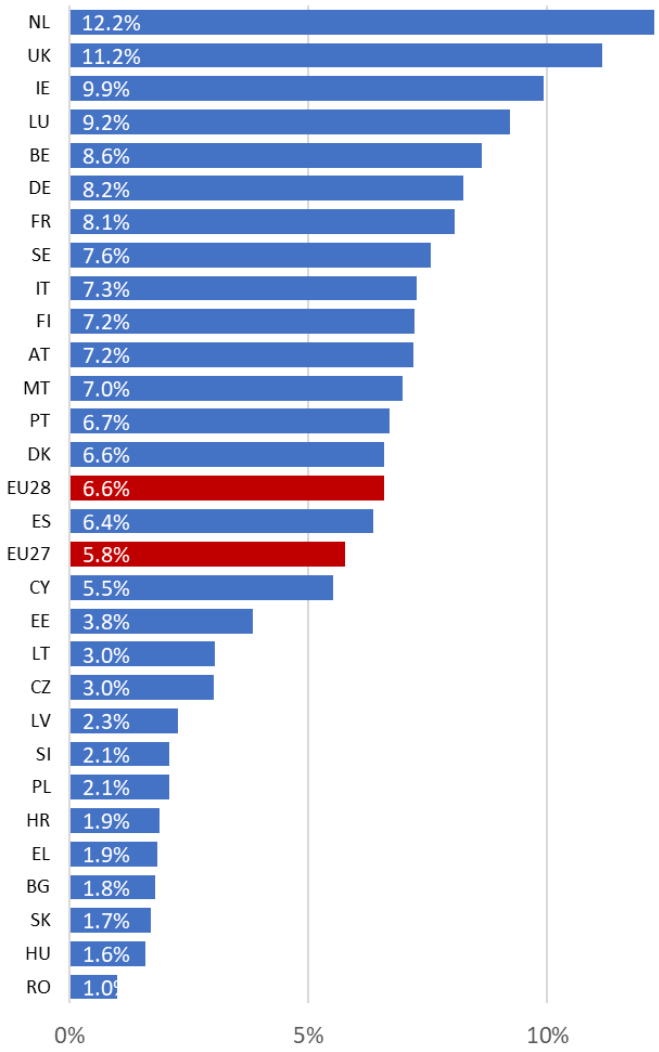



Source: European Data Market Monitoring Tool, IDC 2018

### 5.4.5 Share of Data Users on Total Companies

The penetration rate of data users remains lower than the data suppliers' penetration in Europe. In 2017, our estimates account for a share of 6.6% of data users over the total of companies in the EU28, and 5.8% over the total in the EU27. As a comparison, the share of data users was 6.4% in the EU28 and 5.7% in the EU27 in 2016, thus showing a very modest grow in the incidence of data users over the total number of companies in Europe and the slow dynamic associated with the creation of new companies.

*Figure 16: Data User Companies Share of Total Companies, 2017 (%)*




Source: European Data Market Monitoring Tool, IDC 2018

### 5.4.6 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, as well as Education. 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 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 25 percent to 2025, compared with 1.7 % for those in the smaller size band, they do not add significantly to the total number of data companies.

*Table 31: Data Users forecast 2025 by Member State - Three Scenarios (Units; '000); CAGR 2025-2020 (%)*

| Member State | 2025 Challenge | 2025 Baseline | 2025 High Growth | CAGR 2025/2020 Challenge Scenario (%) | CAGR 2025/2020 Baseline Scenario (%) | CAGR 2025/2020 High Growth Scenario (%) |
|---|---|---|---|---|---|---|
| Austria | 14,600 | 15,050 | 16,350 | 0.7% | 1.3% | 3.0% |
| Belgium | 15,450 | 16,150 | 17,950 | 1.3% | 2.2% | 4.4% |
| Bulgaria | 3,500 | 3,600 | 3,700 | 2.1% | 2.7% | 3.3% |
| Croatia | 1,850 | 1,900 | 1,900 | 1.7% | 2.2% | 2.2% |
| Cyprus | 1,800 | 1,900 | 1,950 | 0.6% | 1.7% | 2.2% |
| Czech Republic | 6,950 | 7,350 | 7,900 | 0.6% | 1.7% | 3.2% |
| Denmark | 8,850 | 9,100 | 10,000 | 1.2% | 1.7% | 3.7% |
| Estonia | 1,900 | 1,900 | 2,050 | 1.1% | 1.1% | 2.6% |
| Finland | 8,350 | 8,650 | 9,000 | 0.4% | 1.1% | 1.9% |
| France | 79,100 | 80,100 | 87,300 | 1.2% | 1.4% | 3.2% |
| Germany | 120,050 | 122,450 | 129,500 | 1.0% | 1.4% | 2.5% |
| Greece | 11,300 | 12,050 | 12,300 | 0.3% | 1.6% | 2.0% |
| Hungary | 5,050 | 5,250 | 5,450 | 1.4% | 22% | 3.0% |
| Ireland | 9,450 | 9,600 | 10,100 | 1.7% | 2.0% | 3.0% |
| Italy | 93,350 | 99,000 | 110,350 | 0.1% | 1.3% | 3.5% |
| Latvia | 1,250 | 1,300 | 1,350 | 0.8% | 1.6% | 2.4% |
| Lithuania | 2,100 | 2,250 | 2,300 | 1.5% | 2.9% | 3.4% |
| Luxembourg | 1,700 | 1,800 | 1,900 | 1.2% | 2.4% | 3.5% |
| Malta | 800 | 800 | 800 | 2.7% | 2.7% | 2.7% |
| Netherlands | 31,100 | 32,750 | 34,400 | 1.5% | 2.6% | 3.6% |
| Poland | 14,300 | 15,350 | 16,700 | 1.0% | 2.5% | 4.2% |
| Portugal | 19,600 | 20,500 | 21,350 | 0.5% | 1.4% | 2.2% |
| Romania | 5,800 | 5,900 | 6,400 | 1.3% | 1.6% | 3.3% |
| Slovakia | 3,200 | 3,300 | 3,700 | 2.3% | 3.0% | 5.4% |
| Slovenia | 1,350 | 1,550 | 1,550 | 0.0% | 2.8% | 2.8% |
| Spain | 76,600 | 81,150 | 87,950 | 0.3% | 1.4% | 3.1% |
| Sweden | 22,900 | 24,450 | 25,050 | 2.1% | 3.4% | 3.9% |
| United Kingdom | 195,850 | 201,500 | 223,950 | 1.5% | 2.1% | 4.2% |
| **EU27** | **562,250** | **585,150** | **629,250** | **0.8%** | **1.6%** | **3.1%** |
| **EU28** | **758,100** | **786,650** | **853,200** | **1.0%** | **1.7%** | **3.4%** |

Source: European Data Market Monitoring Tool, IDC 2018

Table 31 shows the number of data user companies by Member State, and there is little difference in the expected growth by country – as we expect data user companies in all countries to grow at similar rates. The only differences in growth will come from the greater or lesser focus on those industries that are data intensive – specifically Financial services, Retail, Manufacturing.

### 5.4.7 Forecasting Data Users by Industry

No major surprise in the distribution of data users by industry emerges in 2025 under the three scenarios under consideration with respect to the picture portrayed at 2017. Professional services,

Manufacturing, Retail & Wholesale will continue to dominate the scene. In terms of growth, however, the ICT sector will exhibit the most vibrant dynamic through 2025 growing at a faster pace than the EU28.

*Table 32: Data Users Forecast 2025 by Industry - Three Scenarios (Units; '000); CAGR 2025-2020 (%)*

| Industry | 2025 Challenge | 2025 Baseline | 2025 High Growth | CAGR 2025/2020 Challenge Scenario (%) | CAGR 2025/2020 Baseline Scenario (%) | CAGR 2025/2020 High Growth Scenario (%) |
|---|---|---|---|---|---|---|
| Construction | 47,900 | 49,650 | 53,650 | 0.7% | 1.4% | 3.0% |
| Education | 14,800 | 15,250 | 16,700 | 0.8% | 1.4% | 3.2% |
| Financial services | 43,300 | 44,900 | 48,650 | 0.8% | 1.5% | 3.2% |
| Healthcare | 32,500 | 33,550 | 36,150 | 0.8% | 1.5% | 3.0% |
| Information and communications | 85,100 | 88,150 | 96,200 | 1.2% | 1.9% | 3.7% |
| Mining, Manufacturing | 128,850 | 133,900 | 145,150 | 0.7% | 1.5% | 3.2% |
| Professional services | 205,800 | 213,500 | 231,950 | 1.4% | 2.1% | 3.8% |
| Retail and Wholesale | 92,550 | 96,250 | 104,200 | 0.8% | 1.6% | 3.3% |
| Transport and Storage | 91,950 | 95,650 | 103,200 | 0.8% | 1.6% | 3.1% |
| Utilities | 15,350 | 15,850 | 17,350 | 0.9% | 1.5% | 3.4% |
| **EU28** | **758,100** | **786,650** | **853,200** | **1.0%** | **1.7%** | **3.4%** |

Source: European Data Market Monitoring Tool, IDC 2018

Professional services is the largest industry and also shows the highest growth in data user companies as this reflects the data intensive nature of this industry. As well as the volume of data associated with this industry, security and data privacy is important, adding to the growth in the number of data user companies.

### 5.4.8 Forecasting Data Users by Company Size

Again, SMEs are clearly predominant when analysing data users by company size in 2015. Interestingly, however, in 7 years' time, the number of medium to large data user companies is expected to grow at a faster pace and SMEs would account for a lower percentage than in 2017 (and in the years before, according to our European Data Market Monitoring Tool for the period 2013-2015). Indeed, their share is expected to diminish from the 98.8% in 2017 and to be reduced to 94.8% under the High Growth scenario in 2025 (it would be 97.2% in the Challenge Scenario and 95.9% in the Baseline scenario). Much of the Data Market as defined and analysed in this Study needs a degree of expertise and investment which smaller companies are not able to offer. In addition, many smaller organisations tend to be market followers, so look for examples of successful implementation of data transformation before considering making what they consider to be a risky investment. Larger organisations can afford the investment and expertise needed for early adoption of data transformation so the rate of growth for larger companies will be higher than for smaller ones. IDC's 2017 annual survey of technology users supports this with adoption of several data technologies higher among larger enterprises.

*Table 33: Data Users Forecast 2025 by Company Size - Three Scenarios (Units); CAGR 2025-2020 (%)*

| Size Band | 2025 Challenge | 2025 Baseline | 2025 High Growth | CAGR 2025/2020 Challenge Scenario (%) | CAGR 2025/2020 Baseline Scenario (%) | CAGR 2025/2020 High Growth Scenario (%) |
|---|---|---|---|---|---|---|
| 1-249 | 736,900 | 754,250 | 809,200 | 0.7% | 1.2% | 2.6% |
| 250+ | 21,200 | 32,400 | 44,000 | 15.3% | 25.5% | 33.4% |
| EU28 | 758,100 | 786,650 | 853,200 | 1.0% | 1.7% | 3.4% |

Source: European Data Market Monitoring Tool, IDC 2018

### 5.4.9 Forecasting Data Users Average Spend by Company Size

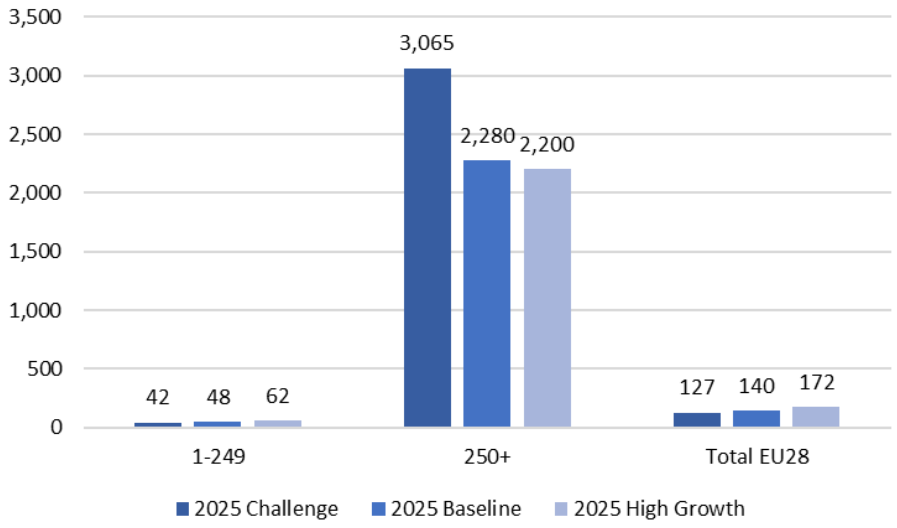
Average spending per company in the Data Market will grow overall, as currently the majority of companies do not invest in data tools and services, but for larger companies this will actually decrease in 2025, as they benefit from decreasing prices for these tools and services. In many cases automation through intelligent agents, cognitive systems, and robotic process automation can reduce overall operating costs, and these tools can replace much more expensive services – which again will lower the cost of investment in data and services.

*Table 34: Data User Companies Forecast 2025 Average Spend by Company Size - Three Scenarios (Units; '000); CAGR 2025-2020 (%)*

| Size Band | 2025 Challenge | 2025 Baseline | 2025 High Growth | CAGR 2025/2020 Challenge Scenario (%) | CAGR 2025/2020 Baseline Scenario (%) | CAGR 2025/2020 High Growth Scenario (%) |
|---|---|---|---|---|---|---|
| 1-249 | 42 | 48 | 62 | 3.9% | 6.6% | 12.0% |
| 250+ | 3,065 | 2,280 | 2,200 | -9.5% | -14.7% | -15.4% |
| Total EU28 | 127 | 140 | 172 | 3.4% | 5.5% | 9.9% |

Source: European Data Market Monitoring Tool, IDC 2018

*Figure 17: Average Spend per Company by Company Size - Three Scenarios (€, '000)*




Source: European Data Market Monitoring Tool, IDC 2018

## 5.5 Analysis of data-driven start-ups in Europe

### 5.5.1 Objectives

In this updated study, the Data Companies' indicator has been designed to explicitly include innovative start-ups and fast-growing SMEs active in the Data Market. The main goal of this analysis is to present an overview of the European landscape by providing additional evidence about the potential number and profile of data-driven start-ups, as well as by tracking the Digital Innovation Hubs (DIHs) with a focus on data science. This is expected to provide valuable insights into the data companies ecosystem, including some statistical evidence on the diffusion of start-ups and an overview of their drivers and barriers in the European Data Market.

To achieve this goal, the study team has designed a suitable methodology, leveraging on a systematic desk research, to map out the landscape of accelerators, incubators and DIHs supporting start-ups, including business angels, business angels' networks, federations of business angels' networks, early stage venture capital funds, business accelerators, business incubators, associates/other early stage market players, universities and scientific parks. More specifically, the analysis has focused on the identification of those accelerators, incubators and DIHs with a strong focus on data-driven innovation, which have been visually mapped in order to provide a direct overview of the landscape. Furthermore, the study team has carried out some targeted interviews with the aim to investigate accelerators, incubators and DIHs' offering portfolio and the type of companies targeted.

### 5.5.2 Overview of the Landscape

In the framework of this updated study, the study team 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 in the European ecosystem. To conduct the research, the study team has relied 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).

In order to gain a deeper understanding of the ecosystem, the study team has carried out two in-depth interviews with:

- Jaime Fernandez Lerga Lopez-Pelegrin from Telefonica Open Future (TOF)[14], the Open Innovation platform of Telefonica[15], which has launched several incubation and acceleration initiatives around the world to promote entrepreneurship and investment in technological projects and which is currently allocating 50% of its investments in big data.
- Paul Czech, International Cooperation and Business Development representative at Know-Center[16], Austria's leading research centre for Data-driven Business and Big Data Analytics, and Coordinator of the European Network of Centres of Excellence in Big Data Research[17], which Know-Center is part of along with other 59 Big Data Centres from 16 Countries.

As regards the potential number of start-ups and innovative SMEs in Europe, the 2017 Funderbeam Start up Investment Report found out that the European ecosystem comprises an average number of 5 start-ups per 100,000 people, with the highest concentration being registered in Iceland, Ireland, Estonia, Finland and the UK, while, according to the Startup Europe Partnership Monitor released in June 2017, around 4,200 fast growing, high-tech companies in the ICT sector can be referred to as "scale-ups", which we can assume represent a key part of the total number of innovative SMEs in the European ecosystem.

Based on data collected from the Startup Heatmap Europe and StartUp Hubs Europe, the most important European start-up hubs in terms of number of companies are located in the UK, Germany and Italy, with London leading the way and being followed by Berlin, Munich and Rome. In terms of investments, however, the situation rather changes: with London maintaining the first position and being followed by Stockholm, Berlin and Paris. These rankings also match with the geographical distribution of start-up associations in Europe investigated through systematic desk-research: in fact, the study team has identified 148 start-up associations in 26 MS, mainly present in the UK, Finland and the Netherlands.

As for the average age of start-ups, in 2016, according to the survey conducted by the European Start up Monitor on a base of 2,365 start-ups, around 29% had less than a year and almost 3 out of 4 had less than 3 years. The survey also highlighted that most start-ups belonged to the IT/software development sector (15%) followed by software as a service (12.2%) and industrial technology/production/hardware (8.3%). No mention, however, was given in relation to the potential number of data-driven start-ups. Instead, an overview of their diffusion has been suggested by the survey realized in 2017 by the European Network of Centres of Excellence in Big Data Research, through which it has been possible to identify 251 data-driven start-ups and 611 SMEs, addressed by a total of 30 research centres belonging to the Network.

### 5.5.3 Supporting Initiatives

Based on its limited and anecdotal evidence, IDC has found out the presence of a lively ecosystem in terms of supporting initiatives in this area, which are providing opportunity of growth for data-

[14] https://www.openfuture.org/en

[15] https://www.telefonica.com/en/home

[16] http://www.know-center.tugraz.at/en/

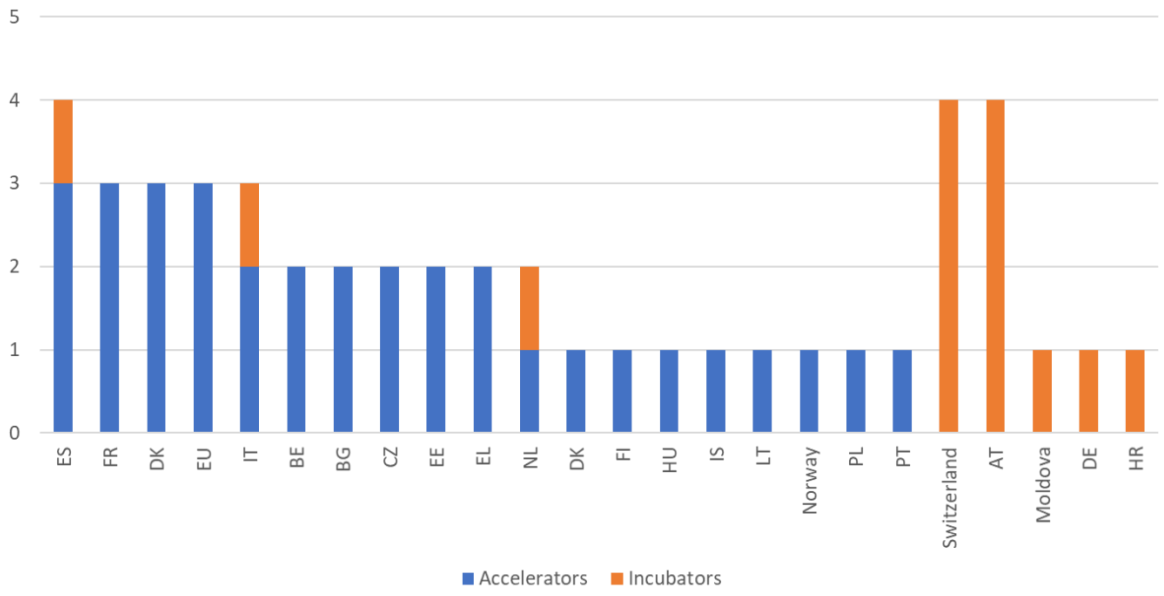
[17] http://www.big-data-network.eu/

driven start-ups and innovative SMEs in the European Data Market. The analysis included in this study has focused on the following three categories: accelerators, incubators and DIHs.

According to the latest European Accelerator Report released in 2016, the number of accelerators in Europe has recorded a growth of 37% compared to 2015. In fact, the report identified 156 accelerators and 193 accelerating programs in 27 countries, with most of them being based in the UK (44), Spain (26), France (23), Germany (17) and Italy (10). Overall, the programs accelerated 3,701 start-ups, 992 of which only in the UK. It is worth mentioning that of the 193 accelerating programs included in the report, 30% described themselves as a combined accelerator, incubator, venture capital fund, and/or angel group and that 54.7% reported an interest in investing in the field of big data analytics during 2017. According to the latest data available on the Startup Europe map, currently registering 169 business accelerators and 124 incubators in Europe, the number has shown an increase over the last year, confirming the growth trend in the evolution of the accelerator industry.

Leveraging on the analysis of interactive maps, such as the aforementioned Startup Europe map, and business incubators and accelerators own websites and directories, the study team conducted in-depth desk-research and found out the presence of approximately 33 accelerators and 14 incubators dealing with data-driven start-ups.

*Figure 18: Accelerators and Incubators of Data-driven Start-ups and SMEs by Country*




Source: IDC desk research, 2018

Among the most interesting initiatives supporting data-driven start-ups with incubation and/or acceleration programmes, it is worth to mention the projects developed in the framework of Horizon 2020: ODINE, Datapitch and the recently announced European Data Incubator (EDI) project.

- **ODINE (Open Data Incubator Europe)**, a 6 month-incubator programme for Open Data entrepreneurs funded with a €7.8m grant, was launched in 2015 by the European Commission and was concluded in April 2017. Its main objective was to attract and fund a group of innovative digital companies in the field of Open data and accelerate their time to market and

chances of success[18]. Apart from contributing to the development of an Open Data ecosystem in Europe, ODINE funded 57 companies, including 31 start-ups born within the programme, whose type of offering can be classified into 3 main groups: pure software, hardware and software, and web-based services. The main services provided by the acceleration programme were training, mentoring, professional advice, networking, access to physical space, industry and infrastructure, direct grant of €100,000 per team, internationalization and communication strategies support[19].

- **Datapitch** is an EU-funded Open Innovation programme launched in October 2017 providing funds and support to data-centred ideas developed by start-ups and SMEs. In the framework of this programme, companies are offered a 6-month business accelerator, up to €100,000 investment, introduction to investors and business partners, peer-networking and support via meetups in major European cities, access to technology and datasets, as well as training and advice by experts[20].
- **The European Data Incubator (EDI)** project, which comprises a budget of more than €7 million, has been launched in January 2018 with the aim to promote R&D and innovation in Big Data. The initiative, led by the University of Deusto in cooperation with 20 organizations from 10 European countries, aims to help companies and start-ups avoid obstacles and develop strategies for the use and exploitation of data, with the main goal to reverse the role that Europe plays in the global market of Big Data, where only two of the top 20 global companies are European. To achieve this goal, EDI will help European SMEs and start-ups by providing them with a high computing and storage infrastructure hosted in the cloud based on open technologies, training and advice in the best-known tools of Big Data Analytics, support to the development of viable business ideas, as well as financing[21].

Apart from incubation and acceleration programmes, other types of supporting initiatives are developed across several Digital Innovation Hubs (DIHs) currently active in the European ecosystem. During its research activity, the study team has been able to identify 506 DIHs in 32 countries (EU28 + Bosnia-Herzegovina, Kosovo, Serbia and Ukraine).

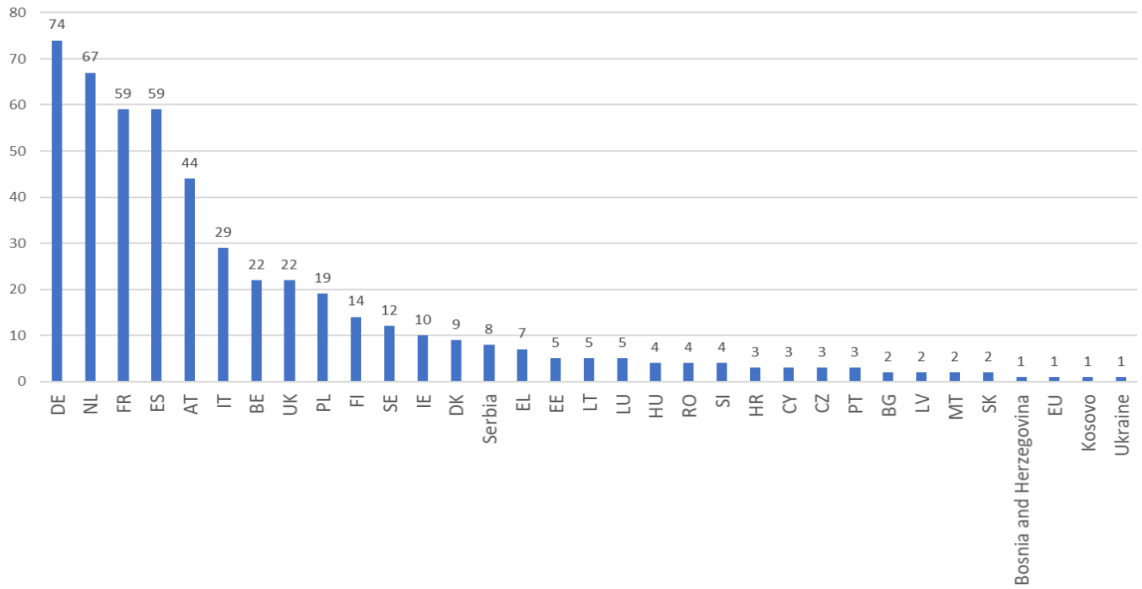
[18] https://opendataincubator.eu/about/

[19] IDC (2017), Impact assessment of ODINE programme.

[20] https://datapitch.eu/about-us/

[21] http://www.zabala.eu/en/news/european-data-incubator-edi-7m%E2%82%AC-boost-big-data-companies

*Figure 19: Digital Innovation Hubs by Country*




Source: Lisbon Council, 2018

## 5.6 Key Findings

**By 2020 there will be nearly 722,000 data user companies** – growing at a compound rate of 2.2%. This is faster than the growth of total companies. The total number of companies is expected to increase at less than 0.2% per year.

**Data users' penetration rates (i.e. number of data user companies on total companies) vary significantly across Member States and present a wider range than the penetration rate exhibited by data professionals (i.e.: number of data professionals on total employment). With the an EU28 average standing at 6.6%, data user' penetration is as low as 1.0% for Romania, and as high as 12.2% for the Netherlands. This relates to the mix of industries associated with each country: those countries with a greater preponderance for data-oriented industries show a higher penetration and will continue to do so. Adoption rates of data technologies are higher in industries such as Professional services, retail, and Financial services.**

**Among the industries, Professional Services is easily the industry with the greatest number of data user companies**, as data is core to a large share of Professional Services activities. Manufacturing follows Professional Services in terms of the number of data user companies, but part of the reason for this is the sheer number of Manufacturing companies among the Member States. Manufacturing is one of the largest industries in the European Union in terms of the number of companies.

**Small and medium companies account for nearly 99% of all data user companies in 2017,** although this is mainly due to the total number of companies in this size band, which is more than 99%. However, average spending by company is disproportionately aligned towards the larger and enterprise companies, with these companies typically spending €5,500,000 per company, compared to a significantly lower €30,000 per company for the small and mid-sized companies. The high spend seen by larger companies will decline out to 2025, and is expected to fall to €2,200,000 by 2025 as economies of scale and wider investment reduce costs for adoption of data tools and services.

# 6. MEASURING DATA COMPANIES REVENUES

## 6.1 Definition

***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. This indicator measures the revenues of the data suppliers identified and classified by Indicator 2, for the products and services specified in our definition of the Data Market.*

The overall value of data revenues is very close, but not identical, to the overall value of the Data Market for the following reasons:

- The value of the Data Market corresponds to the aggregated value of all the data-related products and services bought by European users (demand) including imports from foreign suppliers.
- The value of revenues corresponds to the aggregated value of all the revenues generated by Europe-based enterprises (supply) through the production, distribution, and sale of data-related products and services, including exports outside of the EU.

For the sake of simplicity, it is fair to assume that data industry exports and Data Market imports in the EU are not significant and that tend to balance each other out in a total EU perspective. As in the previous report, this does not preclude the existence of a trade balance surplus or deficit at MS level, which we are unable to estimate for lack of specific evidence.

*Table 35: Data Companies' Revenues: Main Data Sources*

| Data Source | Updated |
|---|---|
| Consensus Forecasts – Consensus economics | Nov-2017 |
| Eurostat chain linked Volumes (GDP) | Dec-2017 |
| IDC Core IT Spending guide 2H2016 | Jul-2017 |
| IDC Worldwide Black Book (standard edition) | Nov-2017 |
| IMF World Economic Outlook | Oct-2017 |

## 6.2 Measuring Data Companies Revenues

Revenues generated by data suppliers have increased by almost 11% in 2017 to reach 52.3 Billion Euros in the EU27 and 68.5 Billion Euro in the EU28 – a constant increase since 2013 according to our initial measurements and the Monitoring Tool. The share of the data suppliers' revenues on the total companies' revenues in the ICT and Professional services sectors has grown as a result and is now stable at 3.2% in both the EU27 and the EU28.

*Table 36: Data Companies Revenues and Growth, 2016-2017-2020 Baseline (€, Million; %)*

| Indicator 3 — Data Companies Revenues and Growth | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| N. | Region | Name | Description | 2016 | 2017 | 2020 Baseline | Growth 2017/2016 | CAGR 2020/2017 |
| 3.1 | EU27 | Total revenues of data companies in | Total revenues of the Data Suppliers calculated by | 47,178 | 52,300 | 69,583 | 10.9% | 10.0% |

| Indicator 3 — Data Companies Revenues and Growth | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | the EU | Indicator 2 | | | | | |
| 3.1 | EU28 | Total revenues of data companies in the EU | Total revenues of the Data Suppliers calculated by Indicator 2 | 61,781 | 68,575 | 89,280 | 11.0% | 9.2% |
| 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% | NA | 6.0% | 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% | 3.2% | NA | 3.4% | NA |

Source: European Data Market Monitoring Tool, IDC 2018

## 6.3 Data Companies Revenues by Member State

The size of the Data Market and of the data-related product and service is the main element continuing to affect the distribution of data revenues across the EU Member States. As a result, the U.K., Germany, France, Italy, the Netherlands and Spain are the Member States with the highest share of data revenues per country - together they account for two thirds (66%) of data revenues in the European Union. The strong software and services bias to these countries' economies is clear from their large share of the market.

Data suppliers' revenues (as seen in Table 37) grow faster than Total ICT spending in the Member States, as organizations invest to reduce cost and grow business through digital transformation. Large Member States in terms of Data Market size and the size of the economy as a whole grow close to the average because of their inertia. The table also points to the fact that high-growth Member States in terms of Data Market-growth have a larger share of suppliers of data tools and software. Figure 20 shows that the U.K. retains its position as the Member State with the largest share of revenue, although Germany, in second place – has higher growth. Not enough to overtake the United Kingdom though.

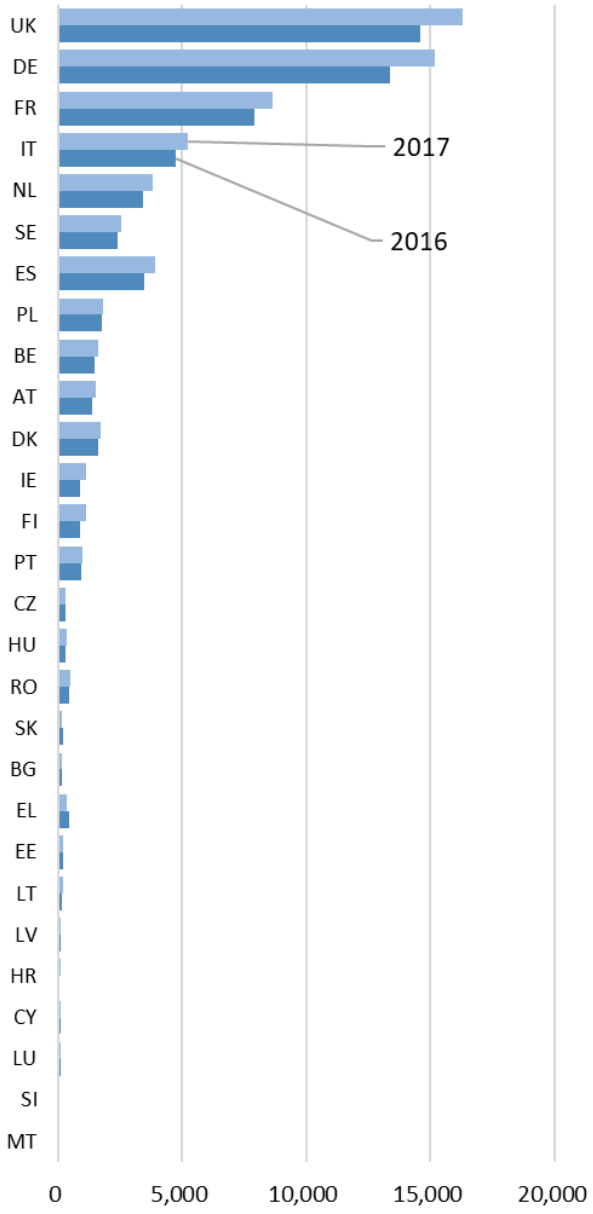
*Table 37: Data Companies'' Revenues by Member State, 2016-2017-2020 Baseline (€, Million)*

| Member State | 2016 | 2017 | 2020 Baseline | Growth 2017/2016 |
|---|---|---|---|---|
| Austria | 1,379 | 1,555 | 1,787 | 12.8% |
| Belgium | 1,477 | 1,615 | 2,305 | 9.3% |
| Bulgaria | 158 | 190 | 300 | 19.8% |
| Croatia | 83 | 102 | 184 | 23.2% |
| Cyprus | 96 | 105 | 182 | 9.2% |
| Czech Republic | 309 | 333 | 637 | 7.8% |
| Denmark | 1,609 | 1,741 | 1,959 | 8.2% |
| Estonia | 199 | 230 | 342 | 15.6% |

| Member State | 2016 | 2017 | 2020 Baseline | Growth 2017/2016 |
|---|---|---|---|---|
| Finland | 918 | 1,127 | 1,501 | 22.7% |
| France | 7,915 | 8,654 | 11,591 | 9.3% |
| Germany | 13,367 | 15,174 | 19,180 | 13.5% |
| Greece | 464 | 362 | 412 | -21.9% |
| Hungary | 315 | 374 | 524 | 18.8% |
| Ireland | 911 | 1,141 | 1,475 | 25.3% |
| Italy | 4,736 | 5,218 | 7,266 | 10.2% |
| Latvia | 118 | 138 | 200 | 16.7% |
| Lithuania | 184 | 207 | 273 | 12.7% |
| Luxembourg | 121 | 124 | 182 | 2.7% |
| Malta | 50 | 49 | 73 | -3.3% |
| Netherlands | 3,424 | 3,831 | 5,324 | 11.9% |
| Poland | 1,768 | 1,842 | 2,334 | 4.2% |
| Portugal | 945 | 983 | 1,138 | 4,1% |
| Romania | 452 | 500 | 627 | 10.6% |
| Slovakia | 200 | 181 | 298 | -9.5% |
| Slovenia | 57 | 73 | 124 | 27.9% |
| Spain | 3,491 | 3,908 | 5,097 | 12.0% |
| Sweden | 2,432 | 2,544 | 4,269 | 4.6% |
| United Kingdom | 14,603 | 16,275 | 19,697 | 11.4% |
| **EU27** | **47,178** | **52,300** | **69,583** | **10.9%** |
| **EU28** | **61,781** | **68,575** | **89,280** | **11,0%** |

Source: European Data Market Monitoring Tool, IDC 2018

*Figure 20: Data Suppliers Revenues by Member State, 2016-2017 (€, Million)*




Source: European Data Market Monitoring Tool, IDC 2018

The key Member States with the highest share of data revenues are the U.K., Germany, France, and Italy, which together account for two thirds (66%) of data revenues in the European Union. The strong software and services bias to these countries' economies is clear from their large share of the market. This dominance will continue as these Members States grow above the average for the EU between 2017 and 2020.

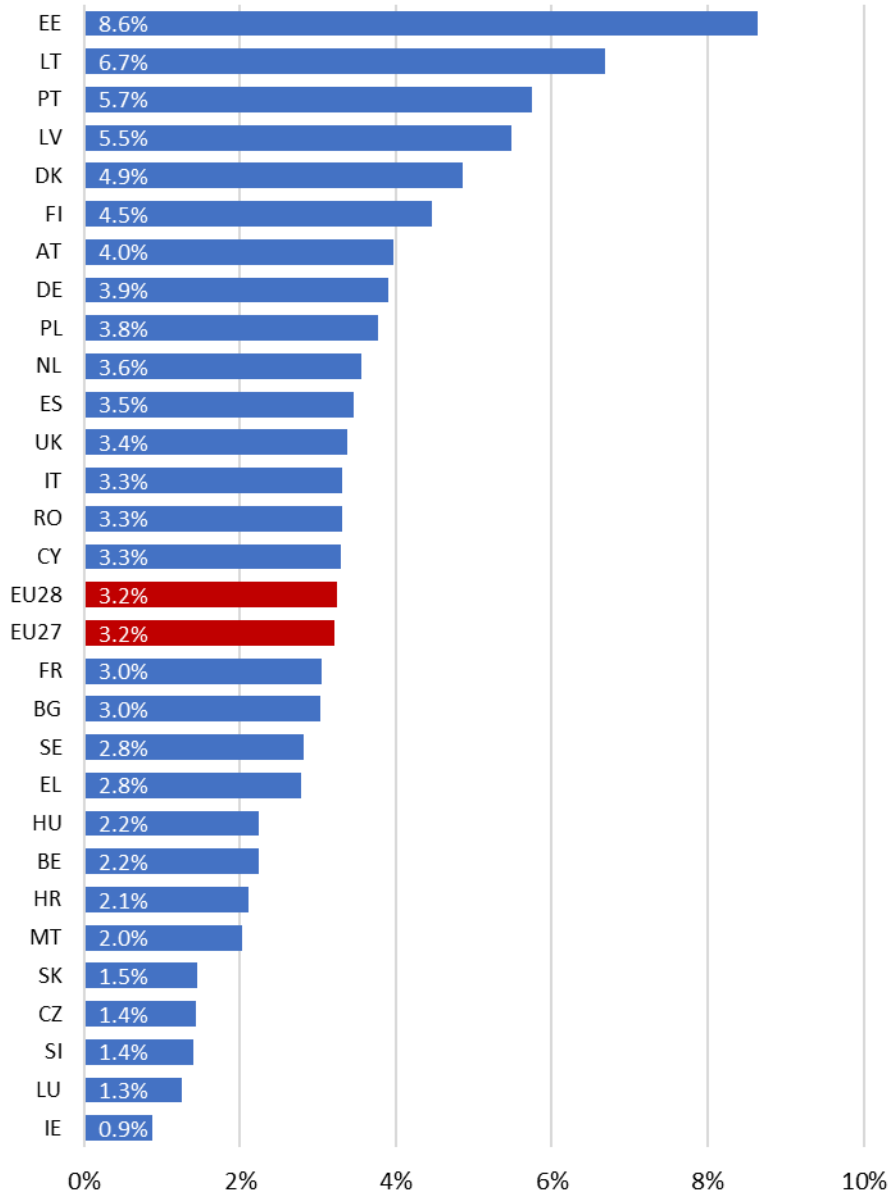
Overall, data companies account for 3.2 percent of total companies' revenues in 2017, but there is a wide spread among the countries as well-known data economies such as Estonia, or the smaller countries with a less well-established services, retail, or manufacturing infrastructure place a higher focus on the Data Market.

## 6.4 Shares of Data Companies' Revenues by Member State

In 2017 we estimate the total data revenues for the EU to be 3.2% of the overall turnover of the data professionals and ICT sectors. Smaller, data-driven and dynamic economies tend to obtain above-average shares of their own data companies' revenues (the Baltics, for example, but also Denmark and Finland), while more traditional economies score primarily below the average (most Central and Eastern European Member States, for instance).

Figure 21 shows the Member States ranked by share of total turnover of data professionals and ICT sectors. The average for all Member States for EU28 and for Member States but excluding the United Kingdom are highlighted in red to provide a reference for the average share.

*Figure 21: Data Companies' Revenues Share of Total Companies' Revenues by Member State, 2017 (%)*




Source: European Data Market Monitoring Tool, IDC 2018

## 6.5 Data Companies Revenues by Company Size

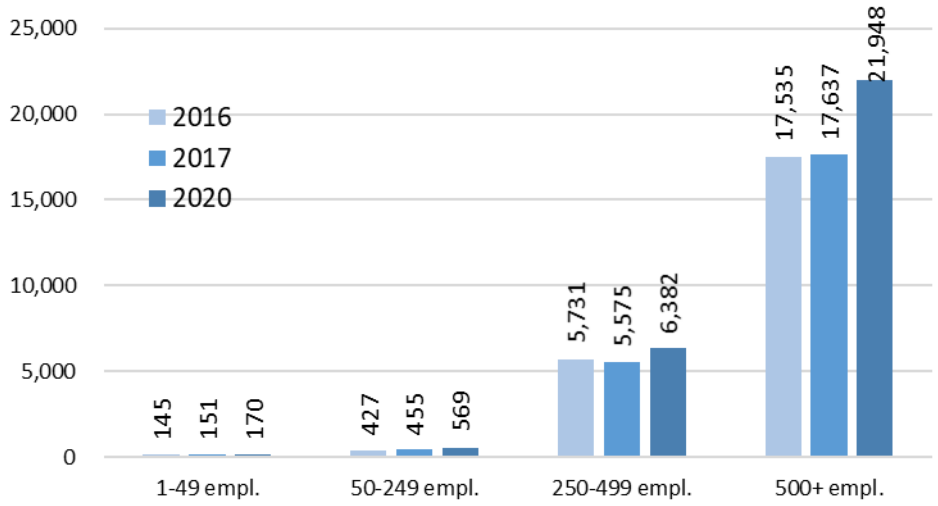
The strength of larger companies in the Data Economy is highlighted in Figure 22, where average revenues for the larger companies – particularly among the 500+ employee companies - shows notable growth – especially to 2020. Small companies don't have the inherent support they need to promote data tools and services.

*Table 38: Data Companies' Revenues by Company Size, 2016-2017-2020 Baseline (€, Million); Growth 2017-2016 (%)*

| Size Band | 2016 | 2017 | 2020 Baseline | Growth 2017/2016 |
|---|---|---|---|---|
| 1-49 empl. | 33,820 | 37,265 | 46,382 | 10.2% |
| 50-249 empl. | 10,908 | 12,296 | 16,658 | 12.7% |
| 250-499 empl. | 15,559 | 17,298 | 23,660 | 11.2% |
| 500+ empl. | 1,493 | 1,716 | 2,580 | 15.0% |
| **EU28** | **61,781** | **68,575** | **89,280** | **11.0%** |
| SMEs share | 72.4% | 72.3% | 70.6% | |

Source: European Data Market Monitoring Tool, IDC 2018

*Figure 22: Average Data Companies' Revenues by Company Size; 2016-2017-2020 Baseline (€, Million)*




Source: European Data Market Monitoring Tool, IDC 2018

Table 38 shows data suppliers' revenues by size class, which reflects the significantly higher number of companies in the smaller size classes. Although Figure 22 shows average revenue by size class is low for the smaller size classes, the total number of companies in the smaller size classes more than compensates for this.

## 6.6 Forecasting the Data Companies Revenues

Data Companies revenues within the EU grow faster than the IT market by some significant amount as these products and services become more mainstream. However, the four major contributing countries (U.K., Germany, France, Italy) lose share between 2017 and 2025 - as their share of EU revenues falls from 66% to 62%. This is not a failing of these countries, but reflects some of the catching up the smaller companies will see as their revenues rise. Over the period 2016-2015 data companies' revenues rise by 8.3% annually, while total IT spending rises are only 1.6%. Figure 24

below shows the different compound growths for the countries, and the four majors are below the average for Europe.

*Table 39: Data Companies' Revenues Forecast 2025 by Member State - Three Scenarios (€, Million)*

| Member State | 2025 Challenge | 2025 Baseline | 2025 High Growth | CAGR 2025/2020 Challenge Scenario (%) | CAGR 2025/2020 Baseline Scenario (%) | CAGR 2025/2020 High Growth Scenario (%) |
|---|---|---|---|---|---|---|
| Austria | 1,948 | 2,349 | 3,162 | 1.7% | 5.6% | 12.1% |
| Belgium | 2,822 | 3,725 | 5,354 | 4.1% | 10.1% | 18.4% |
| Bulgaria | 380 | 558 | 664 | 4.9% | 13.2% | 17.2% |
| Croatia | 182 | 293 | 295 | -0.3% | 9.8% | 9.9% |
| Cyprus | 195 | 256 | 369 | 1.4% | 7.1% | 15.2% |
| Czech Republic | 671 | 844 | 940 | 1.1% | 5.8% | 8.1% |
| Denmark | 1,881 | 2,260 | 3,181 | -0.8% | 2.9% | 10.2% |
| Estonia | 374 | 500 | 700 | 1.8% | 7.9% | 15.4% |
| Finland | 1,310 | 1,519 | 1,826 | -2.7% | 0.2% | 4.0% |
| France | 11,910 | 14,434 | 21,122 | 0.5% | 4.5% | 12.8% |
| Germany | 22,112 | 26,225 | 37,663 | 2.9% | 6.5% | 14.4% |
| Greece | 420 | 521 | 601 | 0.4% | 4.8% | 7.8% |
| Hungary | 597 | 752 | 881 | 2.6% | 7.5% | 10.9% |
| Ireland | 1,758 | 2,233 | 2,847 | 3.6% | 8.6% | 14.1% |
| Italy | 7,181 | 9,536 | 14,795 | -0.2% | 5.6% | 15,3% |
| Latvia | 230 | 307 | 366 | 2.8% | 8.9% | 12,9% |
| Lithuania | 336 | 438 | 519 | 4.2% | 9.9% | 13.7% |
| Luxembourg | 198 | 241 | 327 | 1.7% | 5.8% | 12.5% |
| Malta | 90 | 122 | 165 | 4.5% | 10.9% | 17.8% |
| Netherlands | 6,712 | 8,830 | 12,031 | 4.7% | 10.7% | 17.7% |
| Poland | 3,109 | 3,970 | 5,287 | 5.9% | 11.2% | 17.8% |
| Portugal | 1,133 | 1,444 | 1,898 | -0.1% | 4.9% | 10.8% |
| Romania | 529 | 713 | 1,024 | -3.3% | 2.6% | 10.3% |
| Slovakia | 428 | 579 | 772 | 7.5% | 14.2% | 21.0% |
| Slovenia | 139 | 194 | 254 | 2.3% | 9.4% | 15.4% |
| Spain | 5,226 | 6,873 | 8,416 | 0.5% | 6.2% | 10.5% |
| Sweden | 5,205 | 7,203 | 7,913 | 4.0% | 11.0% | 13.1% |
| United Kingdom | 23,973 | 29,789 | 45,671 | 4.0% | 8.6% | 18,3% |
| **EU27** | **77,077** | **96,921** | **133,372** | **2.1%** | **6.9%** | **13.9%** |
| **EU28** | **101,049** | **126,710** | **179,043** | **2.5%** | **7.3%** | **14.9%** |

Source: European Data Market Monitoring Tool, IDC 2018

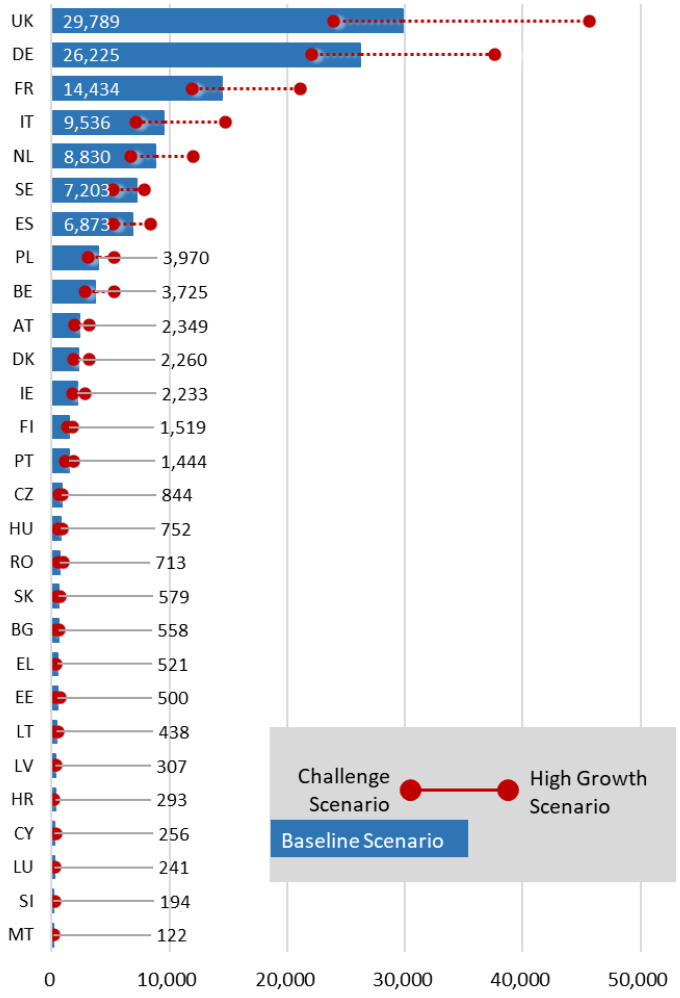
Table 39 shows the expected outcome by Member State for the three different scenarios. The Baseline scenario suggests data revenues will grow by 7.3 percent, which is notably higher than total ICT spending – so data companies' revenues will make a positive contribution to economic growth out to 2025.

Among the Member States included in this table, as usual it is the larger Member States of the United Kingdom, Germany, France, and Italy that have the largest share, but when considering the

overall contribution to growth Germany, France and Italy all grow below average, leaving the United Kingdom to make the most significant contribution to overall growth.

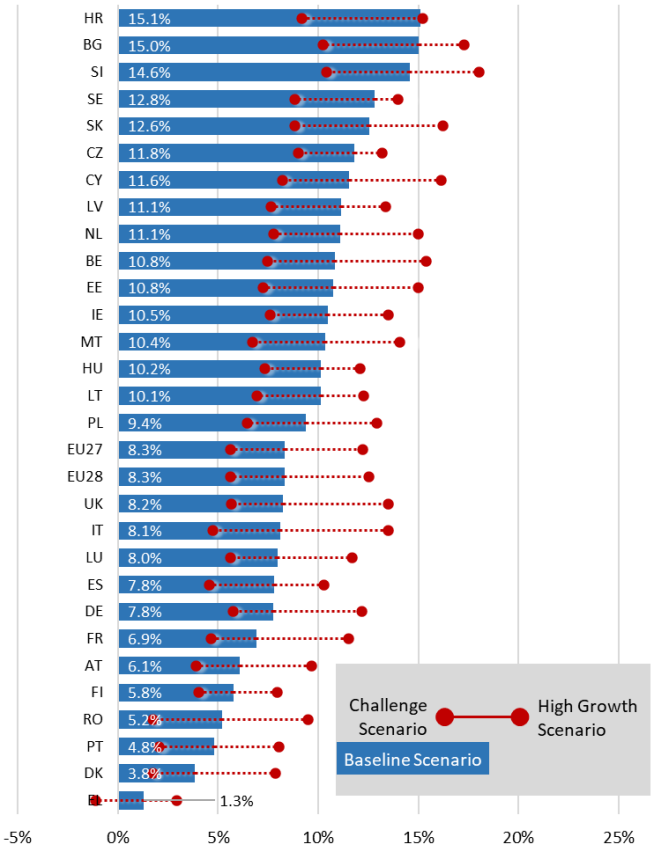
Figure 24 shows the outlook by Member State, and the U.K.'s dominant position is unchanged, but there is a wide range between the Challenge and High Growth scenarios – this reflects the greater uncertainty for the U.K. following political changes in the country. In general, the larger countries have the higher risk for Challenge and High Growth scenarios.

*Figure 23: Data Companies Revenues Forecast by Member State 2025 - Three Scenarios (€, Million)*




Source: European Data Market Monitoring Tool, IDC 2018

*Figure 24: Data Companies Revenues' by Member State CAGR 2025 - Three Scenarios (%)*




Source: European Data Market Monitoring Tool, IDC 2018

## 6.7 Key Findings

**Data companies' revenues will account for 3.2% of total company revenues in 2017**. There is a reasonable range of revenue shares across the Member States as those with a stronger digital presence drive their economies' growth with digital services. Estonia, one of the leading adopters of digital technologies, will have 8.6% of its total company revenues accounted for by the Data Market; Luxembourg has only 1.3%, and Ireland 0.9%.

In the different company size bands, **it is the mid-sized companies that have the greatest revenues**, as the mix of the number of companies and spending are maximized in this middle band. There are many more smaller companies, but they spend little on data tools and services, while the higher spend seen in the larger companies is offset by their low numbers – accounting for less than 1% of all companies. Looking at average company revenues by size band, larger companies show the greatest average spend.

Data 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 8.3%** - easily outpacing the growth of the total ICT market over the same period (expected to be 1.6% from 2016 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.

# 7. MEASURING THE DATA MARKET

## 7.1 Definition

*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 captures 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. Further, the Data Market represents a wider concept than the market of Big Data & Analytics (BDA) as it includes not only the value generated by pure data players developing BDA technologies but also the value created by data-related research, businesses, information and IT services. The digital data exchanged as "products" or "services" in the Data Market refer exclusively to data that 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, image databases). The value of the Data Market is not exactly equal to the aggregated revenues of the European data companies because it 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.

While the Data Market's definition in this Update of the European Data Market Study's report has been left substantially unchanged from the previous results obtained through the European Data Market Study (SMART 2013/0063), the study team has leveraged a wealth of new and updated sources in this edition. The main sources are listed below; additional details are offered in the methodological annex at the end of this document.

*Table 40: The Data Market: Main Data Sources*

| Data Source | Updated |
|---|---|
| Consensus Forecasts – Consensus economics | Nov-2017 |
| Eurostat chain linked Volumes (GDP) | Dec-2017 |
| IDC Core IT Spending guide 2H2016 | Jul-2017 |
| IDC European Vertical Markets survey (2017) | Sep-2017 |
| IDC Worldwide Black Book (standard edition) | Nov-2017 |
| IMF World Economic Outlook | Oct-2017 |
| IT Big Data and Analytics spending Guide 2H2016 | Aug-2017 |

## 7.2 Measuring the Data Market

Table 41 provides an overview of our latest estimates of indicator 4.1 – value and growth of the Data Market. The value of the Data Market in 2017 for both the EU27 and the EU28 is showing a buoyant growth rate of more than 9% year-on-year and is expected to surpass the threshold of 60 billion Euro in 2020 in the EU27 according to the Baseline scenario described in our previous study – 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 measures the Data Market at almost 46.2 billion Euro in 2016.

*Table 41: Measuring the Data Market*

| Indicator 4.1 — Value and Growth of the Data Market (€ Million; %) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| N. | Market | Name | Description | 2016 | 2017 | 2020 Baseline | Growth 2017/2016 | CAGR 2020 Baseline/2017 |
| 4.1 | EU27 | Value of the Data Market | Estimate of the overall value of the Data Market | 46,183 | 50,438 | 60,254 | 9.2% | 6.1% |
| 4.1 | EU28 | Value of the Data Market | Estimate of the overall value of the Data Market | 59,496 | 65,038 | 77,407 | 9.3% | 6.0% |

Source: European Data Market Monitoring Tool, IDC 2018

## 7.3 Data Market by Member State

At Member State level the size of the Data Market remains tightly correlated with the overall economic strength of the Member States, as well as with their national spend on ICT. As in the previous European Data Market Study (SMART 2013/0063), the Data Market continues to exhibit a relatively high level of concentration, with six Member States (the U.K., France, Germany, Italy and Spain – the "Big 5" – plus the Netherlands) accounting for approximately three quarters of the EU28 Data Market amount in 2017.

In terms of year-on-year growth, in 2017 the Data Market dynamic is all the more active in small economies, which are rapidly trying to fill their innovation gap with traditionally innovative Member States by pushing forward their digitization strategies. This is the case of some Central and Eastern European countries such as Slovenia (10.7% year-on-year growth in 2017) and Croatia (10,8%) but also Slovakia (12.1%) and Latvia (10.2%). Among the more traditionally ICT-oriented economies, the Data Market continues to grow significantly in Germany, the Netherlands, and Ireland with year-on-year growth rates well above the EU27 or EU28 average. Table 42 and Figure 25 below provide the details of the Data Market dynamics at Member State level for the years 2016-2017 and 2020 Baseline.

*Table 42: Data Market by Member State - 2016, 2017, 2020 Baseline (€, Million)*

| Member State | 2016 | 2017 | 2020 Baseline |
|---|---|---|---|
| Austria | 1,257 | 1,362 | 1,478 |
| Belgium | 1,540 | 1,658 | 2,274 |
| Bulgaria | 246 | 267 | 332 |
| Croatia | 164 | 182 | 227 |
| Cyprus | 120 | 135 | 173 |
| Czech Republic | 604 | 660 | 705 |
| Denmark | 1,283 | 1,394 | 1,588 |
| Estonia | 199 | 219 | 279 |
| Finland | 981 | 1,060 | 1,107 |
| France | 7,427 | 8,088 | 8,924 |

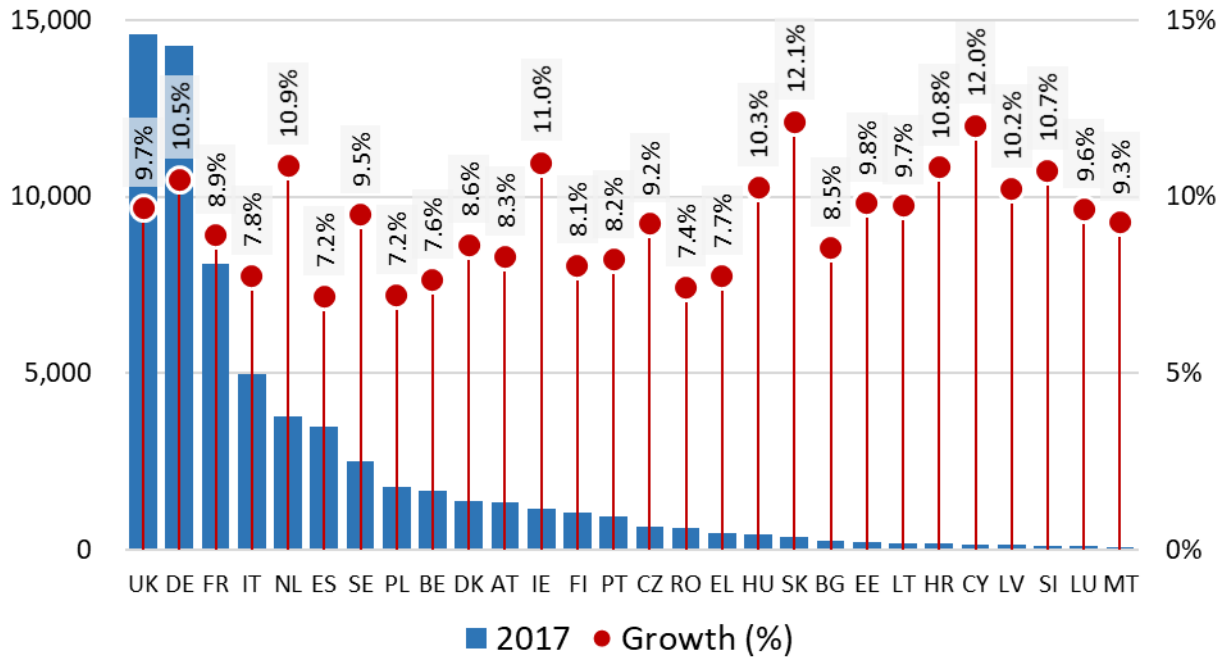
| Member State | 2016 | 2017 | 2020 Baseline |
|---|---|---|---|
| Germany | 12,925 | 14,280 | 16,066 |
| Greece | 430 | 463 | 523 |
| Hungary | 395 | 435 | 532 |
| Ireland | 1,042 | 1,156 | 1,228 |
| Italy | 4,606 | 4,963 | 6,116 |
| Latvia | 118 | 130 | 158 |
| Lithuania | 183 | 200 | 243 |
| Luxembourg | 114 | 125 | 150 |
| Malta | 53 | 58 | 75 |
| Netherlands | 3,395 | 3,764 | 5,432 |
| Poland | 1,660 | 1,780 | 2,081 |
| Portugal | 860 | 930 | 1,080 |
| Romania | 579 | 622 | 620 |
| Slovakia | 322 | 361 | 468 |
| Slovenia | 114 | 127 | 166 |
| Spain | 3,261 | 3,495 | 3,960 |
| Sweden | 2,304 | 2,523 | 4,270 |
| United Kingdom | 13,313 | 14,600 | 17,153 |
| **Total EU27** | **46,183** | **50,438** | **60,254** |
| **Total EU28** | **59,496** | **65,038** | **77,407** |

Source: European Data Market Monitoring Tool, IDC 2018

The results shown in the table above points to a high level of concentration of the Data Market spending among five Member States – United Kingdom, Germany, France, Italy and Spain, although Sweden and the Netherlands overtake Spain by 2020 in terms of its contribution to the overall European Market.

Figure 25 below clarifies this by ranking in order of size the Data Markets of each Member State – Growth for 2017 is shown as the red bubbles - while there does not seem to be a particularly strong relationship between relative size and relative growth, opportunities for higher growth appear to be more affordable by smaller Member States.

*Figure 25: Data Market Size, Growth by Member State, 2017 (€, Million; %)*




Source: European Data Market Monitoring Tool, IDC 2018

## 7.4 Data Market by Industry

Table 43 and figure 26 below display our latest estimates of the Data Market in terms of size and growth by industry in 2016, 2017 and 2020 in the baseline scenario. In line with the results of the previous study, the industries where the Data Market continues to play a dominant role are the sectors characterized by high levels of data-driven innovation manufacturing, Financial services and Professional services in the first place.

In terms of year-on-year growth, the ICT, Financial services and educations are the services experiences the highest growth rates in 2017 and are expected to exhibit healthy growth rates throughout 2020 and beyond.

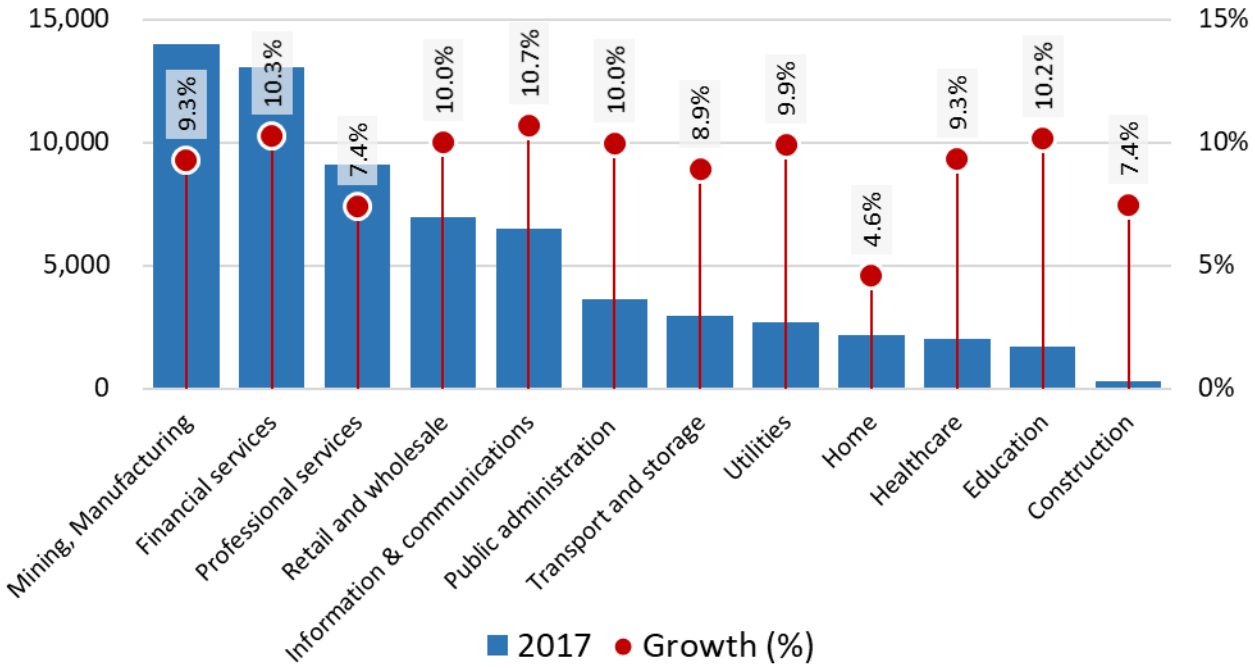
*Table 43: Data Market by Industry, 2016-2017-2020 Baseline (€, Million; %)*

| Industry | 2016 | 2017 | 2020 Baseline |
|---|---|---|---|
| Construction | 275 | 296 | 339 |
| Education | 1,562 | 1,721 | 2,021 |
| Financial services | 11,816 | 13,030 | 15,429 |
| Healthcare | 1,846 | 2,018 | 2,350 |
| Information & communication | 5,865 | 6,492 | 7,818 |
| Mining, Manufacturing | 12,814 | 14,000 | 16,456 |
| Professional services | 8,490 | 9,117 | 11,438 |
| Public administration | 3,281 | 3,608 | 4,244 |

| Industry | 2016 | 2017 | 2020 Baseline |
|---|---|---|---|
| Retail and wholesale | 6,319 | 6,951 | 8,209 |
| Transport and storage | 2,690 | 2,929 | 3,432 |
| Utilities | 2,466 | 2,710 | 3,237 |
| Home | 2,071 | 2,166 | 2,435 |
| Total EU28 | 59,496 | 65,038 | 77,407 |

Source: European Data Market Monitoring Tool, IDC 2018

*Figure 26: Data Market Size, Growth by Industry 2017 (€, Million; %)*




Source: European Data Market Monitoring Tool, IDC 2018

The different industries are listed in Table 43, which shows the biggest Data Market is Manufacturing, which has several sources for data use, from straightforward logistics and resource management, through to the use of the Internet of Things (IoT) which provides a continued data stream for equipment and materials location, and machine activity. However, while Financial services is second in size it shows higher growth than Manufacturing. From this we can say the greatest opportunity is with Financial services, but the industries which need the greatest support are those with the lowest growth – particularly Professional services and Construction.

## 7.5 The Data Market and its Share on Total ICT Spending

The growing importance of the Data Market is confirmed by its comparison with the ICT spending's dynamics, both at Member State and at industry level. For this First Report on Facts and Figures, IDC has reviewed the estimates of the EU27 and EU28 ICT spending and has compared this variable with the size and growth of the updated Data Market estimate.

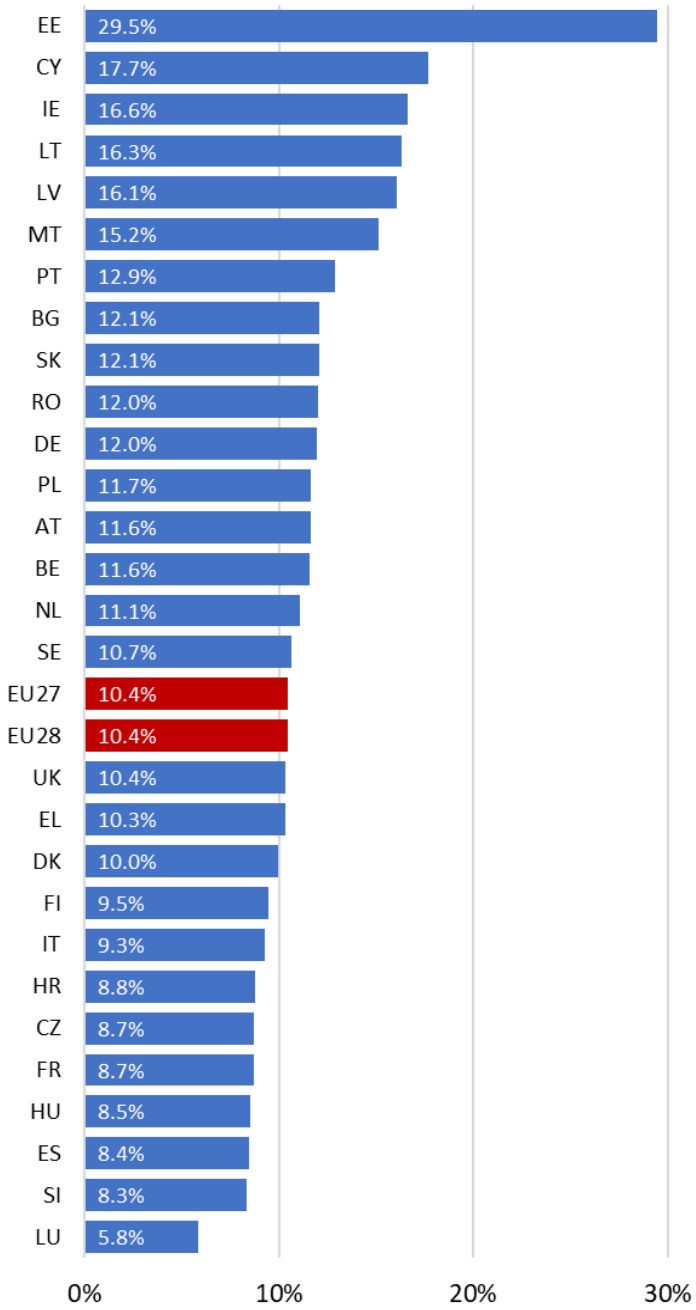
In 2017 the total of the Data Market is expected to represent 10.4% of the overall ICT spending in the EU and has been updated to 11.6% in 2020 according to our previous Baseline scenario. The trend is in constant increase since the beginning of the measurement of this ratio (it was 8.1% in 2013, 8.7% in 2014, 8.8% in 2015, 9.8% in 2016) thus confirming the vitality of the data-related

spending in information and communication technologies if compared to the whole of the ICT spending.

### 7.5.1 Data Market Share on ICT Spending by Member State

Figure 27 provides an overview of the Data Market share on ICT spending by Member States. Not surprisingly, smaller economies tend to present an over-the-average share, either because they still have a relatively small ICT market (this is the case of Cyprus, Latvia or Lithuania, for example), or because they have significantly invested in innovation over the past few years and can now boast a very dynamic market of data-related products and services (such as Estonia, Ireland and, to a lesser extent, Malta, for instance).

*Figure 27: Data Market Share of Total ICT by Member State, 2017 (%)*



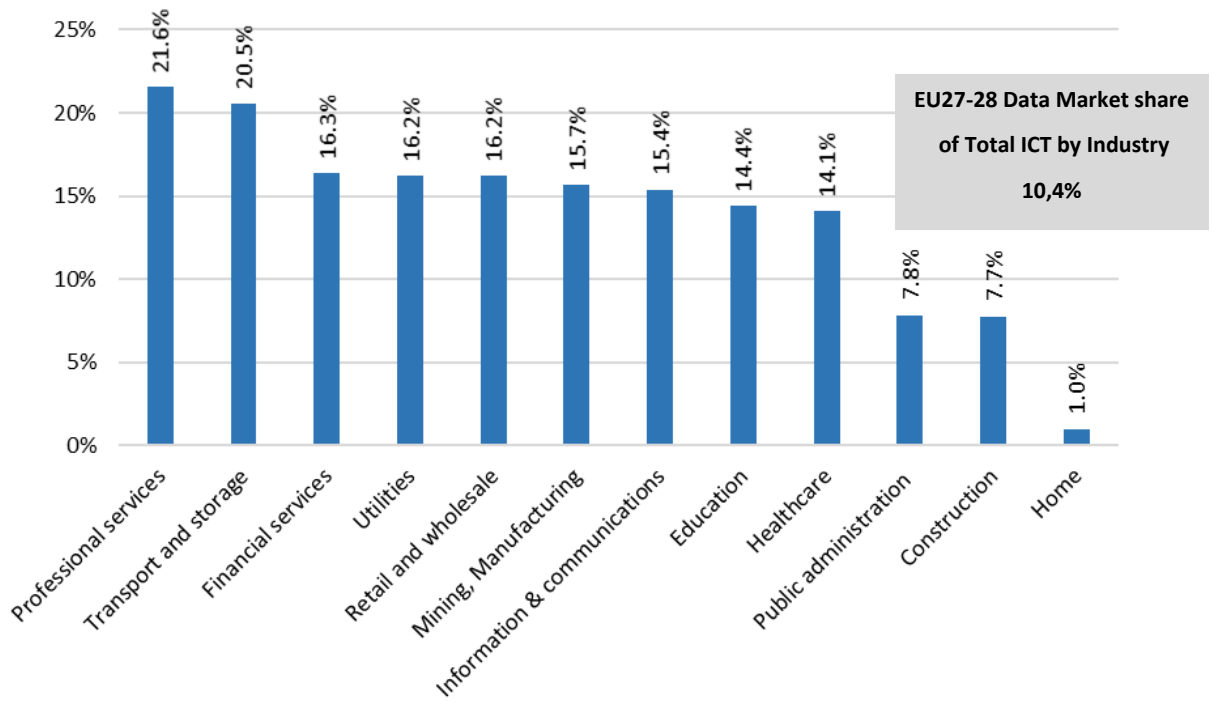

Source: European Data Market Monitoring Tool, IDC 2018

Figure 27 shows the contribution by Member State to total ICT spending which is made by the Data Market. In 2017, just over 10 percent of ICT spending can be attributed to the Data Market, and this will grow as the Data Market will grow faster than Total ICT spending out to 2025. The smaller Member States have a larger share of ICT spending because of their smaller manufacturing and ICT base – using data technology and digital transformation to grow their IT economy. The already established ICT-aware countries such as the United Kingdom, France, Italy, and Spain make a below average contribution. The averages for the EU (EU28) and for the EU without the UK (EU27) are shown as red bars in this chart, to provide a reference.

### 7.5.2 Data Market Share on ICT Spending by Industry

In terms of industry, Professional services, Transport & Storage and Financial services are the three sectors with the highest share of the Data Market over their ICT spending by industry (21.6%, 20.5%, 16.3% in 2017), followed by Retail & Wholesale, Manufacturing and the ICT industries – all with an industry share of Data Market well above the EU28 average over the whole period under consideration.

*Figure 28: Data Market share of Total ICT by Industry, 2017 (%)*




Source: European Data Market Monitoring Tool, IDC 2018

## 7.6 Data Market forecasts at 2025

In 2025 the value of the Data Market under the new High Growth scenario is expected to largely double in size with respect to its 2017 estimates in both the EU27 and the EU28. This will correspond to a considerable CAGR for the period 2020-2025 of 12.7% and 13.6% in the EU27 and the EU28 respectively. According to our new 2025 Baseline scenario, the Data Market will amount to more than 85 billion Euro in the EU27, against 50.4 billion Euro in 2017 (a 7.1% CAGR 2020-2025), while under the Challenge scenario the Data Market will still represent 73.7 billion Euro, growing at a compound annual growth rate of 4.1% 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).

### 7.6.1 Forecasting the Data Market at 2025 by Member State

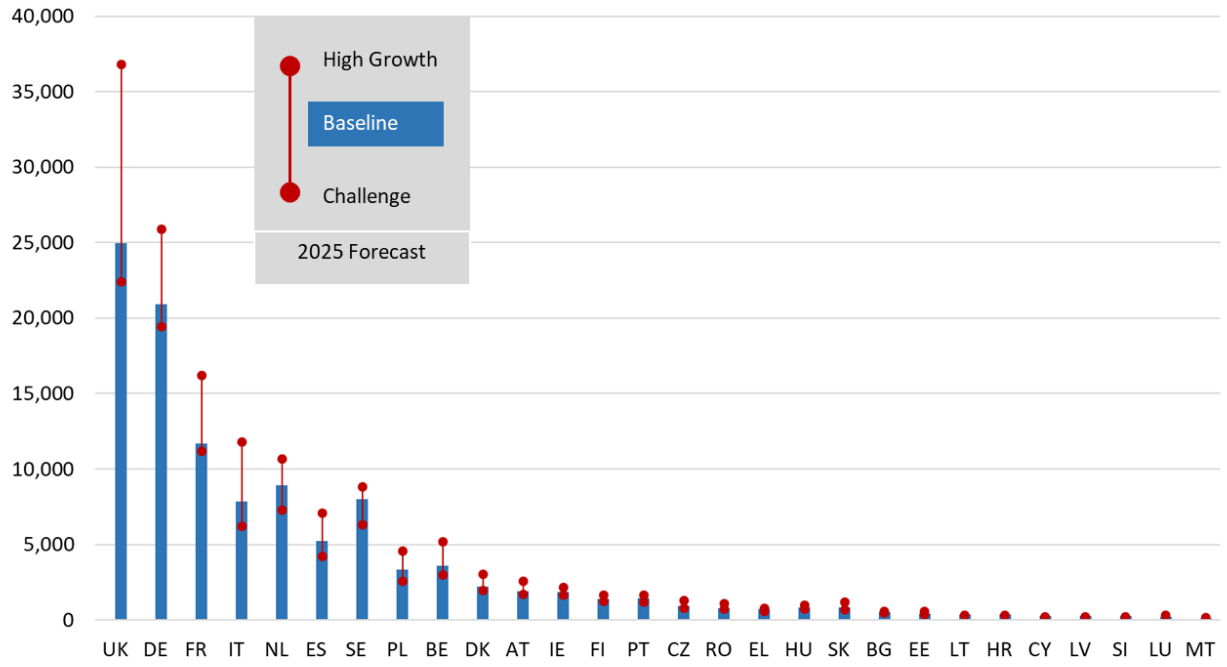
No major surprise appears when looking at the Data Market forecasts at 2025 at the level of Member State with respect to the picture that emerges in the years 2016-2017. The "Big Six" plus the Netherlands and Sweden are clearly dominating the Data Market scene in 2025 under all three scenarios in terms of market size (see Table 44 and Figure 29 below).

*Table 44: Data Market Forecast 2025 by Member State - Three Scenarios (€, Million)*

| Member State | 2025 Challenge | 2025 Baseline | 2025 High Growth | CAGR 2025/2020 Challenge Scenario (%) | CAGR 2025/2020 Baseline Scenario (%) | CAGR 2025/2020 High Growth Scenario (%) |
|---|---|---|---|---|---|---|
| AT | 1,685 | 1,876 | 2,557 | 2.7% | 4.9% | 11,6% |
| BE | 2,986 | 3,600 | 5,192 | 5.6% | 9.6% | 18,0% |
| BG | 472 | 539 | 590 | 7.3% | 10.2% | 12,2% |
| HR | 289 | 337 | 362 | 5.0% | 8.3% | 9,8% |
| CY | 242 | 263 | 276 | 6.9% | 8.7% | 9.7% |
| CZ | 763 | 941 | 1,279 | 1.6% | 6.0% | 12.7% |
| DK | 1,942 | 2,214 | 3,040 | 4.1% | 6.9% | 13.9% |
| EE | 344 | 423 | 577 | 4.3% | 8.7% | 15.7% |
| FI | 1,252 | 1,391 | 1,650 | 2.5% | 4.7% | 8.3% |
| FR | 11,191 | 11,710 | 16,183 | 4.6% | 5.6% | 12.6% |
| DE | 19,430 | 20,900 | 25,907 | 3.9% | 5.4% | 10.0% |
| EL | 570 | 708 | 769 | 1.7% | 6.2% | 8.0% |
| HU | 716 | 822 | 976 | 6.1% | 9.1% | 12.9% |
| IE | 1,653 | 1,825 | 2,163 | 6.1% | 8.3% | 12.0% |
| IT | 6,211 | 7,830 | 11,776 | 0.3% | 5.1% | 14.0% |
| LV | 197 | 236 | 249 | 4.5% | 8.3% | 9.6% |
| LT | 305 | 368 | 382 | 4.6% | 8.6% | 9.5% |
| LU | 206 | 231 | 312 | 6.5% | 9.1% | 15.8% |
| MT | 112 | 124 | 146 | 8.3% | 10.4% | 14.1% |
| NL | 7,272 | 8,900 | 10,674 | 6.0% | 10.4% | 14.5% |
| PL | 2,565 | 3,317 | 4,542 | 4.3% | 9.8% | 16.9% |
| PT | 1,202 | 1,427 | 1,653 | 2.2% | 5.7% | 8.9% |
| RO | 729 | 789 | 1,094 | 3.3% | 5.0% | 12.0% |
| SK | 656 | 842 | 1,161 | 7.0% | 12.4% | 19.9% |
| SI | 202 | 254 | 270 | 4.1% | 8.9% | 10.3% |
| ES | 4,182 | 5,236 | 7,057 | 1.1% | 5.7% | 12.2% |
| SE | 6,302 | 7,974 | 8,795 | 8.1% | 13.3% | 15.5% |
| UK | 22,421 | 24,979 | 36,800 | 5.5% | 7.8% | 16.5% |
| EU27 | 73,677 | 85,078 | 109,633 | 4.1% | 7.1% | 12.7% |
| EU28 | 96,098 | 110,057 | 146,433 | 4.4% | 7.3% | 13.6% |

Source: European Data Market Monitoring Tool, IDC 2018

*Figure 29: European Data Market Forecast 2025 by Member State - Three Scenarios (€, Million)*




*Source: European Data Market Monitoring Tool, IDC 2018*

Figure 29, together with Table 44, show the dominance of the major countries – U.K., Germany, France, and Italy. However, these Member States' position as leaders of the digital economy remains unchallenged. The smaller countries need considerable support if they are to make a notable contribution to the Data Economy, but small increases in growth among the major countries could easily out-match significant changes in the smaller economies. In particular, the high growth impact of the United Kingdom is greater than the combined total of the Data Markets for the bottom 23 Member States.

### 7.6.2 Forecasting the Data Market by Industry

The Data Market dynamic at 2025 will not differ much at industry level too, with Manufacturing, Financial services, Professional services and ICT capturing approximately 70% of the entire Data Market under all three scenarios under consideration. Table 45 ad figure 30 below present our forecasts at 2025 for the value of the Data Market by industry under the three scenarios.

*Table 45: Data Market Forecast 2025 by Industry - Three Scenarios (€, Million)*

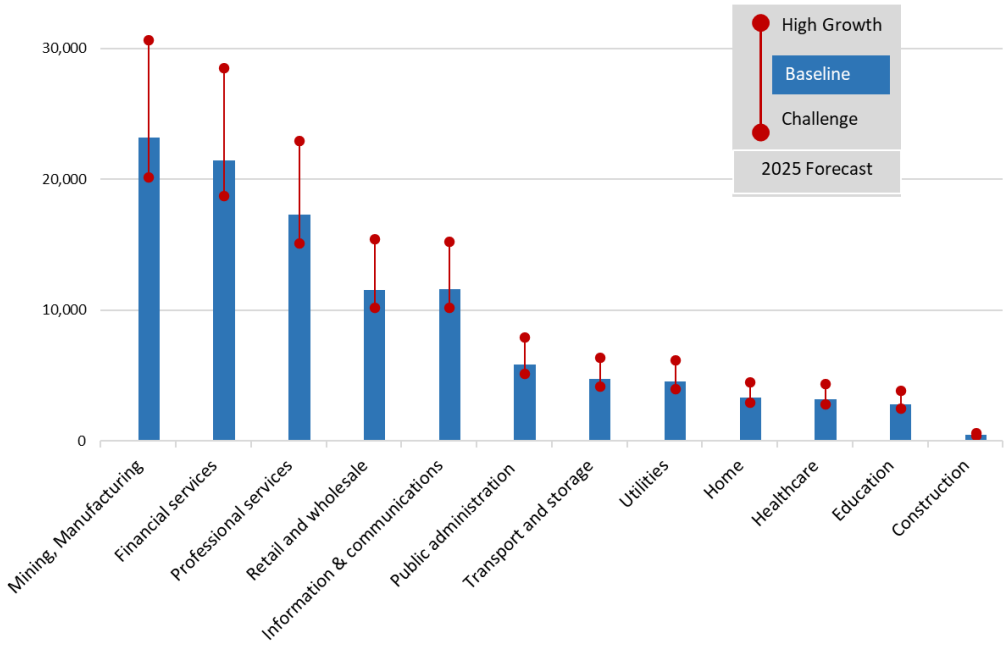
| Industry | 2025 Challenge | 2025 Baseline | 2025 High Grow |
|---|---|---|---|
| Construction | 393 | 448 | 595 |
| Education | 2,480 | 2,818 | 3,850 |
| Financial services | 18,710 | 21,450 | 28,466 |
| Healthcare | 2,787 | 3,207 | 4,365 |
| Information & communications | 10,177 | 11,581 | 15,244 |
| Mining, Manufacturing | 20,117 | 23,178 | 30,658 |
| Professional services | 15,080 | 17,288 | 22,926 |
| Public administration | 5,125 | 5,848 | 7,929 |

| | | | |
|---|---|---|---|
| Retail and wholesale | 10,160 | 11,566 | 15,428 |
| Transport and storage | 4,151 | 4,760 | 6,341 |
| Utilities | 3,984 | 4,568 | 6,157 |
| Home | 2,934 | 3,344 | 4,473 |
| Total EU28 | 96,098 | 110,057 | 146,433 |

Source: European Data Market Monitoring Tool, IDC 2018

Clearly growth in the Data Market will come from Manufacturing, Financial services, Professional services, and ICT – so these are the industries where a small improvement in the market will have a significant overall impact – ahead of the smaller industries such as Utilities, Healthcare, Transport and Home. Support for these key industries to help them transform themselves will have the greatest return for investment. However, many companies are technology followers, and smaller companies are risk averse. The investment required by these companies means they need support and help if they are to contribute meaningfully to the digital economy.

*Figure 30: European Data Market Forecast 2025 by Industry - Three Scenarios (€, Million)*



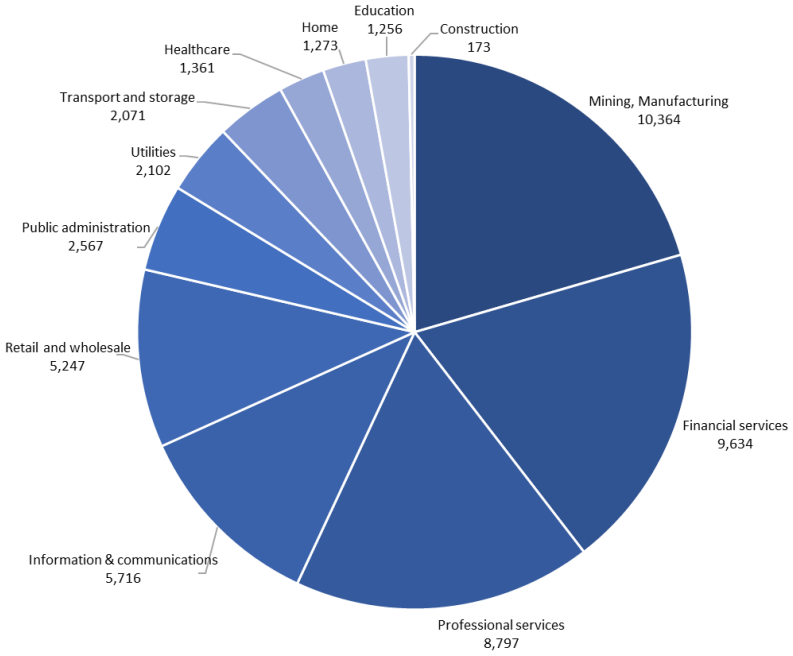

*Source: European Data Market Monitoring Tool, IDC 2018*

A different but insightful way of looking at where growth comes from in 2025 by industry is to take the size of each industry in 2016 and subtract it from the size of the same industry in 2025. This will give the total added Data Market value by industry over the period 2016-2025, i.e. the total increase in spend for that industry over the period under consideration. For the same of simplicity and consistency, the total added market value by industry has been estimated using the 2025 Baseline forecast only, assuming this will be the most likely scenario to happen.

This combination of market growth results with market share results for each of the industries considered in this report provides a powerful means to identify the economic strength of the Data Market by industry in the upcoming years. Not surprisingly, the results are and in line with what was obtained in the previous versions of the European Data Market Study (SMART 2013/0063) - more than three quarters of the added spend will continue to come from only five industries –

Manufacturing, Financial services, Professional services, ICT and Retail & Wholesale, that is the sectors where data-driven innovation is expected to translate into the highest number of use cases and generate the strongest impacts in Europe and elsewhere. Figure 31 below summarises the total added Data Market spend by industry in the period 2016-2025.

*Figure 31: Total Added Data Market Spending by Industry 2016-2025 (€, Million)*




Source: European Data Market Monitoring Tool, IDC 2018

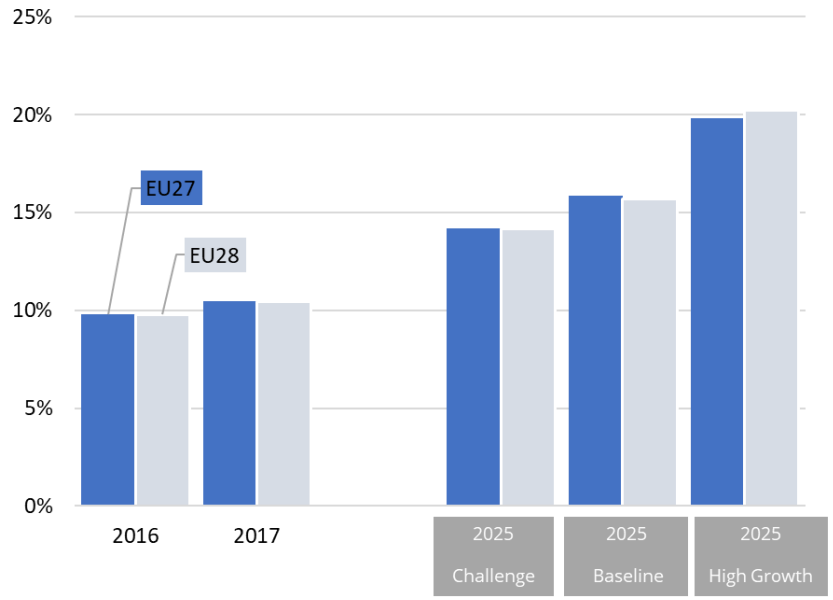
The dominance of growth contribution of the five key industries is clear from Figure 31 - three-quarters of growth in the Data Market is accounted for by these five industries. Small improvements in any of these five will make a notable improvement to the overall Data Market. The large share of the key industries, coupled with their growth mean they are the key contributors to the success of the Data Market in the EU.

### 7.6.3 Forecasting the Data Market Share on Total ICT Spending

By comparing and contrasting our latest estimates of the Data Market forecasts with IDC's forecasts of the ICT spending in 2025 in both the EU27 and EU28, and according to the three scenarios under consideration, we are able to present the expected Data Market share on the total of ICT spending over the forecast period.

In line with our analysis of in 2014, 2015 and 2016, the ratio between the Data Market and the total ICT spending continues to grow in Europe, with a Data Market dynamic that clearly surpasses the ICT developments. As a result, the Data Market share continues to grow and will exceed the threshold of 10% in 2017 in both the EU27 and EU28 (as a comparison it was 8.2% and 8.1% in 2013, 9% and 8.8% in 2014, 9% and 8.8% in 2015 and 9.8% in both the EU27 and the EU28 in 2016). In 2025, we expect the Data Market to represent almost 16% of the ICT spending in the EU27 according to the baseline scenario. This percentage would be only slightly reduced should the adverse scenario materialize (14.2%) and would near the symbolic edge of 20% under the High Growth scenario. Figure 32 summarises the key elements above.

*Figure 32: Data Market share of Total ICT (%)*




Source: European Data Market Monitoring Tool, IDC 2018

## 7.7 Key Findings

**The European Data Market will grow by 9.3% in 2017, and at an average rate of 6.0% out to 2020**. This will make the market valued close to 77.5 €Bn. Each of the Member States shows strong growth, well ahead of the expected growth for the IT market as a whole, which is expected to grow only 2.6% in 2017, and at an annual rate of 1.7% to 2020. The Data Market share of total ICT is 104% for 2017 and will grow out to 2025.

**Among the Member States, Estonia stands out the Member State as having a Data Market with the largest share of ICT spending**, but Estonia has embraced digital economics for some time, so its position is not surprising. Aside from Estonia, the spread of share of ICT spending taken by the Data Market is fairly narrow across the Member States with the largest share being 17 percent, and the smallest is 5.8%

**The larger industries account for the largest share of the Data Market – as there is the greatest number of companies within these industries, but adoption rates of Data Technology tend to be high in finance, Professional services, and in retail, which, together with the size of these industries makes them 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 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 2016 to 2025, with more than half coming from the three main industries of Manufacturing, Finance, and Professional services.**

**The Data Market will continue to out-grow the total ICT market, with its share of this market rising from close to 10% in 2016, to more than 15% by 2025.**

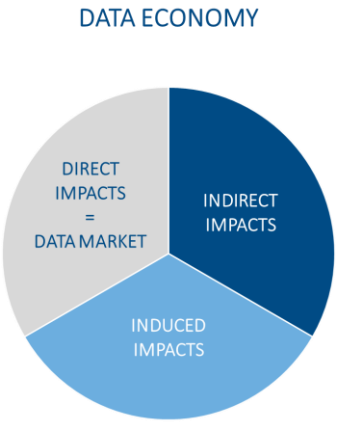
# 8. MEASURING THE DATA ECONOMY

## 8.1 Definition

*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 represents a wider concept than the Data Market. The latter is the market where digital products or data-enabled services are exchanged, generating innovation, improving production processes, and allowing for enhanced customer experience and understanding; the Data Market is therefore a business-oriented concept. The Data Economy instead captures a wider reality as it apprehends the value and wealth generated in the economy as a whole (not just across businesses) by the exploitation of data. The Data Economy does include the Data Market in the form of direct impacts on the economy and it is further composed of indirect and induced impacts.

*Figure 33: The Data 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. 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 will coincide with the value of the Data Market.
- **The indirect impacts:** 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 (Richardson, 1985):
    - **The backward indirect impacts**: such impacts represent the business growth generated in those businesses that provide inputs to the data suppliers. In order to produce and deliver data products and services, the data companies need inputs from other stakeholders. Revenues from sales to data suppliers are the backward indirect impacts.
    - **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):

  - Optimising production and delivery processes: data-driven processes (data-driven production);
  - Improving marketing by providing targeted advertisements and personalised marketing practices (data-driven marketing);
  - Improving existing organisation and management practices (data-driven organisation).

- **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 will support economic activity in various industries such as retail, consumer goods, banks, entertainment, etc.

Our Data Economy estimation 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 (monetary) terms, they are not economic impacts as defined above as they do not induce an increase in the business activities and a consequent growth in GDP.

## 8.2 Measuring the Data Economy

### 8.2.1 Data Economy Results

Table 46 provides an overview of the latest updates of indicators 4.2 – value and growth of the Data Economy. The value of the Data Economy for the EU28 has been estimated to exceed the threshold of EUR 300 billion in 2016, and the estimated 2017 growth rate of 12% highlights a faster growth, considering the 7% for 2016. The share of overall impacts on GDP is expected to grow from 2.2% in 2016 to 2.4% in 2017. Results for EU27 and EU28 are similar, but it is worth highlighting the difference in the 2020/2017 CAGR and in 2017 growth rates for the two regions. EU27 rates are higher than EU28, highlighting uncertainty related to Brexit is affecting the U.K. which is showing lower than average growth rates and 2020/2017 CAGR.

*Table 46: Measuring the Data Economy*

| Indicator 4.2 — Value and Growth of the Data Economy (€ Million; %) | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| N. | Market | Name | Description | 2016 | 2017 | 2020 Baseline | CAGR 2017/2016 | CAGR 2020/2017 | Impact on GDP 2016 | Impact on GDP 2017 |
| 4.2 | EU28 | Value of the Data Economy | Value of the Data Market plus estimate of direct, indirect and induced impacts on the economy | 305,977 | 335,618 | 452,190 | 9.7% | 10.4% | 2.2% | 2.4% |
| 4.2 | EU27 | Value of the Data Economy | Value of the Data Market plus estimate of direct, indirect and induced impacts on the economy | 243,205 | 267,006 | 365,761 | 9.8% | 11.1% | 2.1% | 2.2% |

Source: European Data Market Monitoring Tool, IDC 2018

## 8.3 Data Economy by Member States

### 8.3.1 Value of Data Economy and Data Economy on GDP

Table 47 provides an overview of the latest updates of the Data Economy from a country perspective. There are countries expected to grow at a faster pace than other countries in 2017, Eastern and Central European countries in particular (Bulgaria, Cyprus, Estonia, Latvia, Lithuania, and Slovakia among the fastest ones). Indeed, these countries are more receptive than mature economies to the digital transformation process that helps them in filling the gap in innovation. This strength is reflected in a faster growth, while the size of the Data Economy is considerably lower than big countries, considering that in 2017 the six Member States (the U.K., France, Germany, Italy and Spain – the "Big 5" – plus the Netherlands) represent 80% of the total Data Economy for EU28.

In terms of impact on GDP, the picture is variegated. Estonia, the U.K., Germany, the Netherlands, and Denmark stand out with shares well above the EU27 or EU28 average.

*Table 47*: Data Economy by Member State, 2016-2017-2020 Baseline (€ Million), Growth Rates, CAGR (%)

| Member State | 2016 | 2017 | 2020 Baseline | GR 2017/2016 | CAGR 2020/2017 | Impact on GDP 2016 | Impact on GDP 2017 |
|---|---|---|---|---|---|---|---|
| AT | 7,693 | 8,270 | 10,396 | 7.5% | 7.9% | 2.4% | 2.5% |
| BE | 7,814 | 8,752 | 12,851 | 12.0% | 13.7% | 2.0% | 2.2% |
| BG | 654 | 791 | 1,291 | 21.0% | 17.7% | 1.5% | 1.8% |
| HR | 462 | 553 | 881 | 19.8% | 16.8% | 1.0% | 1.2% |
| CY | 318 | 395 | 683 | 23.9% | 20.1% | 1.7% | 2.1% |
| CZ | 2,147 | 2,457 | 3,332 | 14.4% | 10.7% | 1.2% | 1.4% |
| DK | 6,714 | 7,317 | 9,531 | 9.0% | 9.2% | 2.6% | 2.7% |
| EE | 533 | 655 | 1,097 | 23.1% | 18.7% | 3.0% | 3.5% |
| FI | 4,495 | 4,979 | 6,518 | 10.8% | 9.4% | 2.4% | 2.5% |
| FR | 38,557 | 42,405 | 57,001 | 10.0% | 10.4% | 1.8% | 2.0% |
| DE | 78,323 | 84,122 | 110,669 | 7.4% | 9.6% | 2.7% | 2.9% |
| EL | 1,336 | 1,486 | 2,138 | 11.2% | 12.9% | 0.7% | 0.8% |
| HU | 1,113 | 1,343 | 2,158 | 20.6% | 17.1% | 1.0% | 1.2% |
| IE | 5,535 | 6,495 | 8,324 | 17.3% | 8.6% | 2.2% | 2.5% |

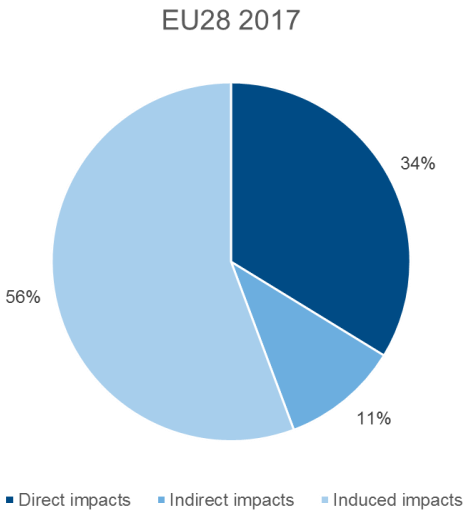
| | | | | | | | |
|---|---|---|---|---|---|---|---|
| IT | 28,632 | 29,683 | 37,925 | 3.7% | 8.5% | 1.8% | 1.9% |
| LV | 320 | 394 | 632 | 22.9% | 17.1% | 1.5% | 1.8% |
| LT | 521 | 638 | 1,021 | 22.5% | 16.9% | 1.5% | 1.8% |
| LU | 386 | 460 | 673 | 19.4% | 13.5% | 0.8% | 0.9% |
| MT | 193 | 233 | 351 | 21.1% | 14.6% | 2.2% | 2.5% |
| NL | 16,928 | 19,412 | 29,284 | 14.7% | 14.7% | 2.5% | 2.8% |
| PL | 5,096 | 5,412 | 5,685 | 6.2% | 1.7% | 1.2% | 1.2% |
| PT | 2,865 | 3,247 | 4,679 | 13.3% | 13.0% | 1.6% | 1.8% |
| RO | 1,682 | 2,010 | 2,704 | 19.5% | 10.4% | 1.1% | 1.3% |
| SK | 875 | 1,103 | 1,888 | 26.1% | 19.6% | 1.1% | 1.4% |
| SI | 356 | 422 | 686 | 18.5% | 17.6% | 0.9% | 1.1% |
| ES | 21,231 | 23,921 | 34,743 | 12.7% | 13.2% | 1.9% | 2.1% |
| SE | 8,425 | 10,051 | 18,617 | 19.3% | 22.8% | 2.0% | 2.3% |
| UK | 62,772 | 68,612 | 86,429 | 9.3% | 8.0% | 3.0% | 3.2% |
| **EU27** | **243,205** | **267,006** | **365,761** | **9.8%** | **11.1%** | **2.1%** | **2.2%** |
| **EU28** | **305,977** | **335,618** | **452,190** | **9.7%** | **10.4%** | **2.2%** | **2.4%** |

Source: European Data Market Monitoring Tool, IDC 2018

### 8.3.2 Value of Data Economy by Type of Impact in EU28 in 2017

The pie chart below provides an overview of the distribution of Data Economy across type of impacts in 2017 for EU27 and EU28. The most important message of this figure is the highest value of indirect impacts in the composition of Data Economy. It reveals how in 2017 the indirect impact, forward impacts in particular, (involving all data users and their exchange with data suppliers), has the highest share in the Data Economy, in particular compared to the induced impacts. Indeed, induced impacts are still not mature and not widespread, and the period is too short to measure the full potential of data products and services. The creation and distribution of data value takes time to realise, especially at the level of induced impacts. In the next paragraph will see the detail by Member State, and in that perspective, differences emerge in the relative weights between indirect and induced impacts.

*Figure 34: Data Economy by Type of Impact, EU28, 2017 (%)*



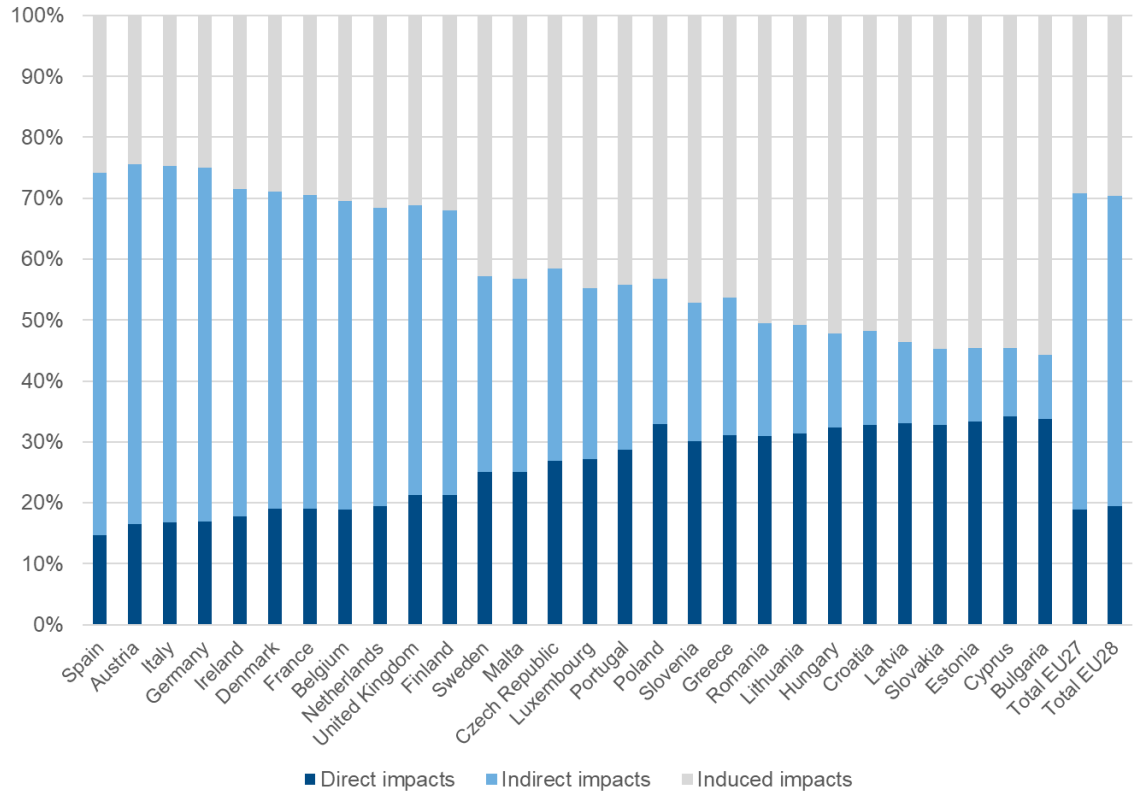

Source: European Data Market Monitoring Tool, IDC 2018

### 8.3.3 Value of Data Economy by Type of Impact by Member State 2017

While overall induced impacts are lower than indirect impacts, at Member State level differences emerge. In particular, there is a considerable variance between Western European economies and countries from Central and Eastern Europe, in which induced impacts have a higher share than indirect impacts. What really changes the incidence of impacts is a widespread diffusion of data-driven innovation throughout the whole economic system. Central and Eastern Europe have significantly invested over the past few years to fill the innovation gap (such as Estonia, for example), and they are facing the digital transformation process benefits.

It is possible to identify some clusters of countries with specific reference to those in which there is a predominance of indirect impacts and a low share of induced impacts. These countries are mainly from Western Europe (Spain, Austria, Italy, Germany, Ireland, Denmark, France, Belgium, the Netherlands, the U.K., and Finland), and represent more mature markets in which data suppliers are already generating huge indirect impacts along the supply chain. Another cluster is composed by Eastern economies (such as Romania, Lithuania, Hungary, Croatia, Latvia, Slovakia, Cyprus, Bulgaria). The ratio of total impacts on GDP in these countries is lower than average, and this lower penetration of the Data Economy is confirmed by the fact that the data suppliers are generating their own revenues, but have a low impact on the supply chain. A full-scale trickle-down effect has not started yet and it is not expected to start in the near future. Estonia can be defined as an outlier, with a more developed Data Market, and a bigger Data Economy than similar countries (which has a higher than average Data Economy and GDP ratio).

*Figure 35: Data Economy by Type of Impact by Member State, EU28, 2017 (%)*




Source: European Data Market Monitoring Tool, IDC 2018

## 8.4 Forecasting the Data Economy

We forecast the impacts at 2025 according to the three scenarios presented in Chapter 3: Challenge, Baseline and High Growth.

- In the Baseline scenario, Europe will exhibit a more solid recovery with GDP trends similar to the pre-crisis years and an increasing ICT spending. Such positive macro-economic trends will support investments for innovation and, therefore, the adoption of data products and services. The Baseline scenario highlights the impacts of data products and services with a positive economic trend. As a consequence, also induced effects will grow.
- In the Challenge scenario, Europe is going to recover slowly, and the GDP growth and the ICT spending are going to increase gradually: this will limit the innovation propensity of companies and, therefore, the adoption and diffusion of new products and services.
- The High Growth scenario foresees macro-economic trends similar to the Baseline scenario. Nevertheless, the ICT technology push will support companies in daring investments helping cost savings and new benefits. Users will keep demanding data products and services, thanks to an awareness effect about the achievable benefits of the data innovation. As a consequence, in the High Growth scenario, the demand of data products and services from companies will grow, but induced impacts and overall effects in the deepest level of the economy will grow more than in the Baseline scenario.

### 8.4.1 Data Economy Forecast: Scenarios and impact on GDP

Table 48 represents the Data Economy forecast at 2025. When comparing the Baseline scenario at 2025 to the Data Market estimated in 2017, we can see that the value will double along the period (from around EUR 336 billion in 2017 to nearly EUR 670 billion in 2025), with a CAGR 2017/2020 of 9.0% for EU28 and 9.3% per EU27, still remaining high when considering the CAGR 2020/2025 (8.2% EU28 and 8.3% EU27). The CAGR 2020/2025 in the EU28 is higher (17.2%) than the CAGR 2017/2025 (14.6%), thus highlighting an acceleration of growth after 2020, that will make the Data Economy for EU28 surpass the threshold of EUR 1 trillion, and accounting for 6% of the GDP at 2025. Similarly, the Challenge scenario will see a slowdown of the economic effects, from 4.3% CAGR 2017/2025 to 0.8% CAGR 2020/2025, with the Data Economy being well below EUR 500 billion.

*Table 48: Data Economy Forecast at 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 |
|---|---|---|---|---|---|---|---|---|
| 4.2 | Value of the Data Economy EU28 | Value of the Data Market plus estimate of direct, indirect and induced impacts on the economy | 470,402 | 669,197 | 1,001,073 | 3.0% | 4.2% | 6.0% |
| 4.2 | Value of the Data Economy EU27 | Value of the Data Market plus estimate of direct, indirect and induced impacts on the economy | 376,520 | 544,241 | 769,505 | 2.9% | 4.0% | 5.4% |

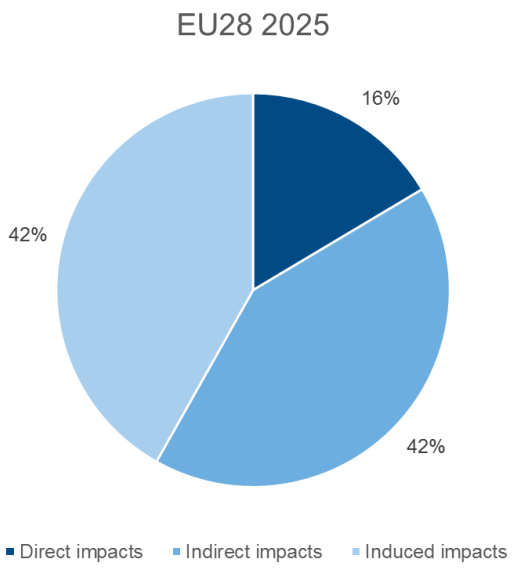
Source: European Data Market Monitoring Tool, IDC 2018

### 8.4.2 Value of Data Economy by Type of Impact in EU28 in 2025

The pie chart below provides an overview of the distribution of Data Economy by type of impacts in 2025 for EU27 and EU28. It is worth highlighting how the composition of impacts changes along time, from 2017 to 2025, in favour of induced impacts, this 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 around 30% in 2017. Indirect impacts in turn will lose around 10% of share, but still in 2025 accounting for a very high percentage (42%). With respect to 2017, 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 36: Data Economy by Type of Impact, EU28, 2025 (%)*




Source: European Data Market Monitoring Tool, IDC 2018

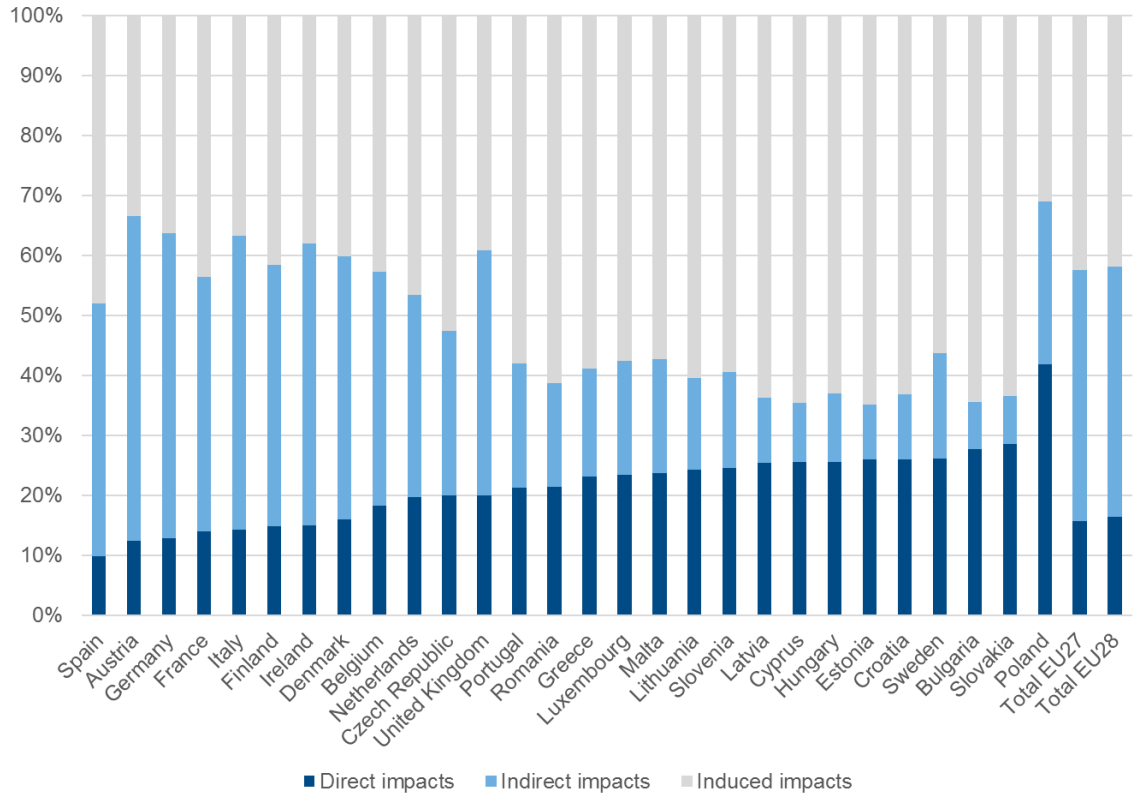
### 8.4.3 Value of Data Economy by Type of Impact by Member State in 2025

The figure below displays the impact distribution by Member State for the EU28 in 2025. With respect to 2017 in almost all countries induced impacts will exceed indirect impacts. As the penetration rates increase, also the distribution of impacts at an induced level increase. In Central and Eastern Europe, induced impacts will increase their share compared to indirect impacts, while in Western Europe indirect and induced impacts will be more balanced, showing a strong and consolidated data supply market in 2025.

Also in 2025 we have countries clusters similar to 2017. The differences between clusters emerge when comparing how impacts are distributed across the three types. In particular, bigger and more mature economies, as the first group of countries show (Spain, Austria, Germany, France, Italy, Finland, Ireland, Denmark, Belgium, the Netherlands, and the U.K.), there is a well-balanced distribution between indirect and induced impacts. In particular, in these countries, the increase of induced impacts in mainly due to a reduction of indirect impacts, rather than a reduction of direct impacts, highlighting how the indirect impacts (forward) are spreading effects in the whole economy. Inside the second group of countries, mainly composed by Central and Eastern European economies, the increase in induced impacts is instead due more to a reduction in direct impacts than

indirect, highlighting some degree of impenetrability across vertical markets and a supply chain that struggles to develop.

*Figure 37: Data Economy by Type of Impact by Member State, EU28, 2025 (%)*




Source: European Data Market Monitoring Tool, IDC 2018

### 8.4.4 Data Economy – Distribution of Impacts – Scenarios Comparison

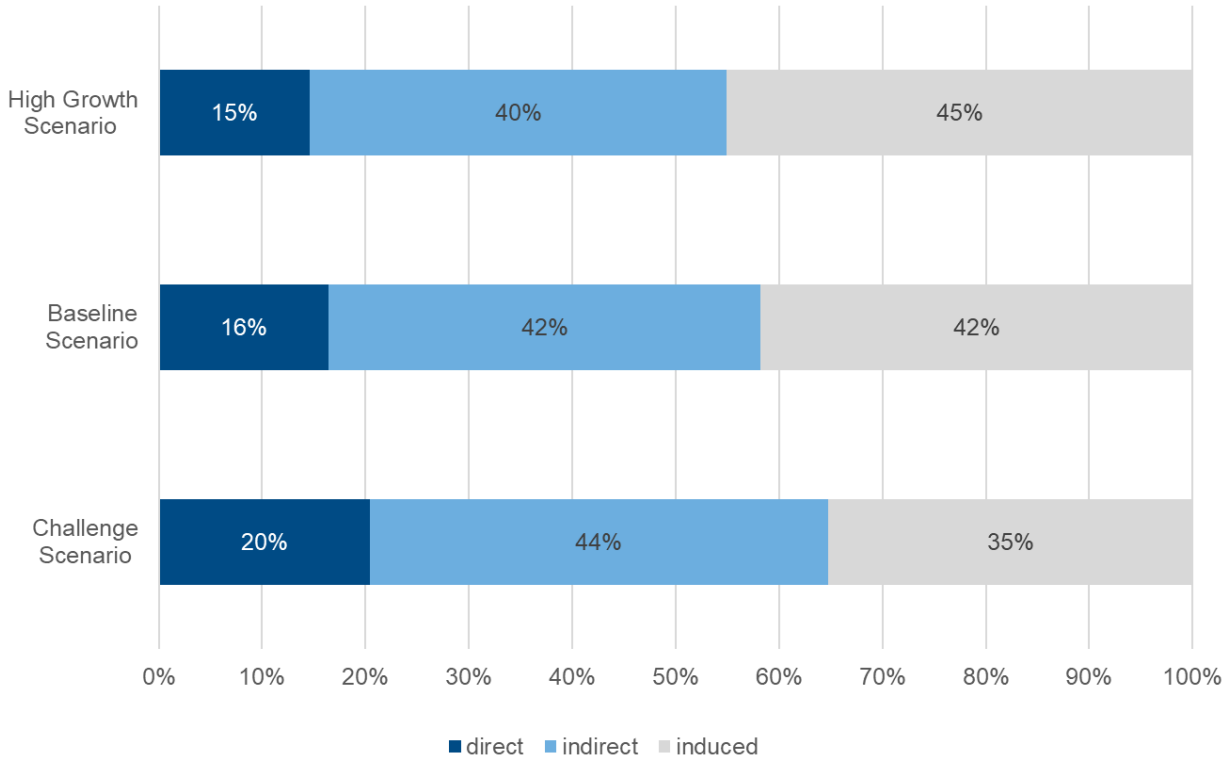
Impacts are unevenly distributed across the different scenarios in EU28 in 2025, showing how the underlying assumptions will affect the three potential development paths. Overall, forward indirect impacts are confirmed to be the most important, despite losing share across the period 2017-2025. These impacts are all the impacts deployed through the user industries. In 2025, the forward indirect impacts will still be supported by a considerable penetration rate of data as well as by increasing benefits due to gains in efficiency and competitiveness in the user companies. Nevertheless, in the previous estimations we have confirmed a conservative estimate of the induced impact because data professionals do not correspond to net job creation and a share of the data professionals are people already working in the companies, dealing with management or ICT. In the estimation through 2025, we considered a five-year period in which the induced effects would increase, as the acceleration of the overall effects in the economy.

- In the Baseline scenario, indirect and induced impacts will be well balanced, accounting for 42% both. With respect to 2017, in which forward are the most relevant impacts, in 2025 induced impacts will increase. We consider the ongoing Brexit process, and we will measure the EU27 total impacts, that will account for 4.2% of GDP.
- In the High Growth scenario, the penetration of data into the user industry is still high, but we considered a slower penetration rate of data services and products into the user

industry, in favour of higher induced impacts, that will be higher than in the other two scenarios. Overall impacts in 2025 for EU27 on GDP will be 5.4%.

- In the Challenge scenario, forward indirect impacts are the most important. The penetration rate of data services and products into the user industry is higher and the induced effects are considerably less relevant than in the other two scenarios. Overall impacts in 2025 for EU27 on GDP will be 2.9%.

*Figure 38: Data Economy by Type of Impact by Member State, EU28, 2017 (%)*




Source: European Data Market Monitoring Tool, IDC 2018

## 8.5 Data Economy by Industry

### 8.5.1 Value of Data Economy by Industry

So far, the Data Economy model has considered the impacts by country. In this Data Economy update, we included an additional view of the resulting impacts: the view by vertical market. One of IDC's strengths is to identify vertical markets' trends, specific features, and differences across each industry. This capability has been applied here to diversify the impacts in terms of industries, with the aim to answer the question "how and to what extent vertical markets convey data value effects to the economy?". IDC believes this approach adds value to the estimation process and to the results.

By leveraging IDC past and existing research on IT spending and on innovation accelerators (cognitive, big data, IoT, among others) adoption across vertical markets we identified industries in which the role of data is predominant, and with a key role in redistributing the value of data to the rest of the economy. In general, companies across all vertical markets are embarking on projects

that revolve around data. This includes data generated by business processes (for analysing where optimisation is needed) and customers (for gathering insight and adopting a more customer-tailored approach). Projects based on Big Data and analytics, IoT, and artificial intelligence will have increased uptake, as companies will increase their focus on data to generate a positive impact on business processes and customer experience. Overall IDC forecasts that in 2017, finance and manufacturing will be the biggest IT spenders, and retail and Professional services will generate the fastest CAGRs in the period 2017-2021. Industries with a higher value of the total impacts are identified in Financial services, Manufacturing, and Professional services, similarly to what found on Data Market estimation in the previous chapter. IT trends for these markets highlight that:

- In Financial services, as security and data breach prevention remain strong priorities, next-generation security applications and blockchain technologies will gain traction. Moreover, the digital distribution channels continue to expand and innovative technologies play a crucial role in allowing banks their quality of services and security.
- In Manufacturing, digital transformation is restructuring the way manufacturers operate their factories, introducing cost-efficient and automated approaches to production. In the Industry 4.0 era, factories are "smart" thanks to interoperability between machines, devices, sensors, and people, enabled by IoT, cloud, cognitive computing, and robotics. The Industry 4.0 model requires cyber-physical systems that monitor processes to be autonomous and decentralised decision makers, simplifying processes and introducing extreme automation.
- In Professional services regulatory compliance is strong and will lead to IT investments as the industry remains one of the most regulated. Cloud is now becoming mainstream while big data technologies are increasingly widespread. Especially larger companies such as big accounting firms are already making significant investments in adopting tools and methodologies surrounding audit data analytics. Data-intensive sub-industries such as the audit profession may indeed gain strongly from analytics.

The table below shows the distribution of total impacts across industries for EU28 in 2017, 2020, and 2025.

*Table 49:* Data Economy by Industry, EU28, 2017, 2020, 2025 (€, Million)

| Industry | Total Impacts 2017 | Total Impacts 2020 | Total Impacts 2025 |
|---|---|---|---|
| Financial services | 64,900 | 86,907 | 135,601 |
| Construction | 1,397 | 1,899 | 2,989 |
| Mining, Manufacturing | 70,635 | 94,854 | 124,641 |
| Education | 8,533 | 11,525 | 18,030 |
| Healthcare | 9,780 | 13,301 | 20,706 |
| Public administration | 17,893 | 24,558 | 38,266 |
| Information & communications | 35,434 | 48,000 | 75,395 |
| Professional services | 53,365 | 71,903 | 113,408 |
| Retail and wholesale | 34,623 | 46,456 | 72,714 |
| Transport and storage | 14,577 | 19,625 | 24,233 |
| Utilities | 14,116 | 19,237 | 25,054 |
| Home | 10,365 | 13,923 | 18,161 |
| **EU28** | **335,618** | **452,190** | **669,197** |

Source: European Data Market Monitoring Tool, IDC 2018

## 8.6 Key Findings

**The indirect impacts will be the key impacts in all the scenarios**, while induced impacts at the same time will start increasing, as the effects in the economy, through job and revenue creation as additional impacts will start to be visible.

**In the baseline scenario the indirect and induced impacts will be more balanced than in 2017 and 2020. The overall impacts for EU27 will be 4.0% of the GDP**. The indirect impacts will still play a fundamental role, as the user industries will consolidate the quantitative benefits from the use of data, but more than that, these benefits will go beyond the users and will translate in higher induced effects, generating jobs and revenues beyond the data companies itself.

**With respect to the previous data delivered, from 2020 and 2025 there will be a change in the share of indirect impacts across the three scenarios**. In particular, in the high growth scenario the penetration rate of data products and services into the user industry will be lower, and replaced by higher induced effects. The positive conditions under the high growth scenario will lead the overall impacts exceed 1 trillion Euros in 2025.

**The estimation of impacts by industry revealed how financial sector, manufacturing, and Professional services are the markets in which the impacts are most relevant**. Indeed, as also shown in the Data Market estimation, these industries make a significant usage of data-related technologies, their forward and backward impacts are high, and can convey effects at an induced level more than other industries thanks to several trends, such as the diffusion of IoT and Cloud computing, the digital transformation processes, as well as the usage of mobile and social technologies.

# 9. MEASURING DATA PROFESSIONALS SKILLS GAP

## 9.1 Definition

*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.*

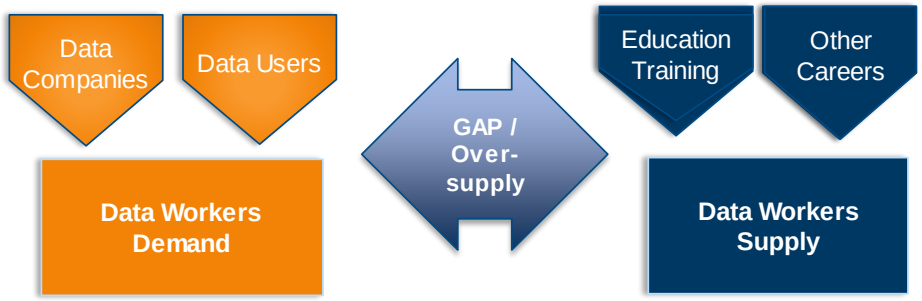
This indicator is based on a model balancing the main sources of data skills (from the education system and re-training and other carriers) with the estimated demand (by all data companies).

More specifically, we use the following definitions:

- **The supply of data professionals** is equal to the data skills supply stock (the sum of employed data professionals and the unemployed ones).
- **The demand of data professionals** is the sum of existing and open positions, that is the number of currently employed data professionals (indicator 1 in this study) plus the unfilled vacancies.
- **The indicator measures the difference between total demand and supply**; if demand is higher than supply there is a data skills gap (excess demand). If supply is higher than demand there is over supply and potentially unemployment.

The demand is the number of data professionals estimated in this study, while supply is estimated aggregating the number of graduates in the relevant disciplines and the level of inflows from other careers or upskilling. The model takes into account the inflows and outflows of the data skills market such as retirements, unemployment and vacancies.

*Figure 39: The Data Skills Demand-Supply Balance Model*




Source: European Data Market Monitoring Tool, IDC 2016

## 9.2 Main Results

This 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 2017, given the strong increase of demand of data professionals from 2016 (+7%), the estimated gap grew by 10% reaching approximately 449,000 unfilled positions, corresponding to 6% of total demand (see Table 50). By 2020 we expect the gap to expand to 699,000 unfilled positions in the EU27, corresponding to 8% of total demand. However, 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 a data skills gap of 6% of demand as in 2017 is not a very large one, while the 8% gap foreseen for 2020 identifies stress in the market.

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 10% of demand in the Challenge one, to 11% in the Baseline, to 19% in the High Growth (EU27), as shown in the following Table 50. The slightly lower dynamic of data skills demand in the UK (due to slightly slower forecast data market growth) means that the gap incidence is smaller in the EU28 than in the EU27. The absolute size of the data skills gap is relevant, potentially reaching 1.1 million unfilled positions in 2025 in the EU 27 Baseline scenario but up to over 2 million in the EU 27 High Growth scenario. In the Challenge scenario the data skills gap is forecast at 904 thousand 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 main trends explaining the forecast data skills gaps presented in the following table are the following:

- Gradual increase of the number of graduates in the period 2020-2025, driven by increasing awareness of the market potential and by the policies promoting STEM (Science, Technology, Engineering and Mathematics) and data skills education. The initiatives promoting STEM were launched in several countries since 2015 and their effects are starting to be seen in the forecast period. For example, Ireland has already seen an increase of 54% of level 8 graduates in the STEM field from 2012 to 2018[22];
- Increase of inflows in the data skills market from other careers, upskilling and re-training initiatives, due to higher attractiveness of the ICT career. This trend affects the Baseline and even more so the High Growth scenario, because of the higher growth dynamics of the data market;
- However, the growth of supply both from graduates and inflows is not expected to keep up with the growth of demand in all 3 scenarios;
- In the Challenge scenario, lower inflows to the data skills market from other careers, upskilling and re-training initiatives, because of the slower development of the Data Market;
- Based on our definition, data professionals comprise a wide portfolio of skills and may come from a wide range of disciplines, from STEM to social science, business and law. There is a high number of graduates with the right skills who may step in if the data professionals' career grows in attractiveness because of strong demand. The dynamics of supply of relevant graduates are projected to be positive across Europe, but more could be done to accelerate them.

[22] Presentation at the workshop "Forecasting ICT high-level skills demand", on the ICT skills Action Plan 2014-2018, by Higher Education Policy and Research Department of Education and Skills, Irish Government, 2018

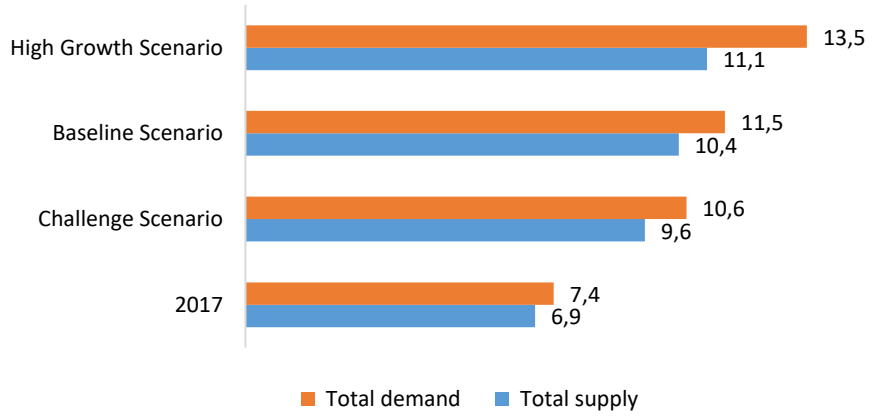
*Table 50: Indicator 6 - Data Professionals Skills Gap in the EU, 2016–2025 - Three scenarios*

| Indicator 6 — Data Professionals Skills gap | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| N. | Name | Description | | Actual | | Baseline | Baseline Scenario | | Challenge Scenario | | High Growth Scenario | |
| | | | | 2016 | 2017 | 2020 | 2025 | 17-25 CAGR | 2025 | 17-25 CAGR | 2025 | 17-25 CAGR |
| 6.1 | Data professionals skills gap | Gap between demand and supply of data professionals, N, 000s | EU27 | 343 | 378 | 603 | 1007 | 13% | 904 | 11.5% | 2,116 | 24% |
| | | | EU28 | 428 | 449 | 699 | 1111 | 12% | 992 | 10.4% | 2,393 | 23.3% |
| 6.2 | | Gap between demand and supply of data professionals, % | EU27 | 6.2% | 6% | 9% | 11% | | 10.4% | | 19% | |
| | | | EU28 | 6.2% | 6% | 8% | 10% | | 9.4% | | 18% | |

Source: European Data Market Monitoring Tool, IDC 2018

The supply of data professionals is less elastic than the demand, responding less quickly to the evolution of economic conditions. When the higher education system adds new skills, for example, the new graduates will enter the labour market with an average time lag of 5 years. This means that the skills gap tends to be higher in the scenarios where demand grows faster. However, the market responds to increased demand also through additional training, learning on the job and inflows from other careers, and this tends to counteract the potential gap. Internal mobility between the MS is also likely to be a factor, but the actual size of these internal EU inflows for specific categories of professionals is very difficult to estimate. On the other hand, we see that the growth of total demand has strong variations between the 3 scenarios, which require anticipated action by policy makers to counteract the potential data skills gap.

*Figure 40: EU 28, Total EU Supply-Demand of Data Professionals, 2017-2025, Three Scenarios (€, Million)*




Note: demand includes potential vacancies

Source: European Data Market Monitoring Tool, IDC 2018

## 9.3 Measurement Approach

The measurement of this indicator is based on a model combining the separate estimates and forecasts of the demand and supply of data professionals with data skills. The definition and measurement approach of data professionals is explained in Chapter 4 and in the Methodological Annex included in this report. The total demand of data professionals for the years 2016 and 2017 includes the data professionals currently employed (as calculated by the Indicator 1) plus estimated vacancies per year (currently unfilled positions). The forecast demand of data professionals to 2020-2025 under the 3 scenarios presented in Chapter 4 is the total potential demand (it incorporates future potential vacancies). The supply is estimated aggregating the data professionals currently employed, the unemployment rates, and the inflows and outflows to the data worker market (including retirements, changed careers, upskilling and so on). The methodology was developed for the forecast of the supply-demand balance of ICT skills (e-skills) in the EU by IDC and empirica 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 2017. 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. The model requires a high number of assumptions and the results are more reliable for the total EU28 than for the individual Member States, because we lack sufficient evidence about the actual/potential mobility of data professionals and markets demand-supply mismatches at national level.

To update the measurement of the data skills gap, the study team has adjusted the model 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.
- Data-driven innovation for growth and well-being, www.oecd.org/sti/ieconomy/data-driven-innovation.htm, 2016.
- ILOSTAT (International Labour Organization) Statistics and Databases (2016).
- EUROSTAT Tertiary Education Statistics (Last update: December 2015)[23].

## 9.4 The Indicator by Member State

The charts below present the data skills gap estimates by scenario for the major countries, highlighting the different dynamics of the Data Market as well as of the national labour markets. More specifically:

- The U.K., one of the leading Data Markets in Europe, because of Brexit consequences reducing economic growth potential moves from above EU average to below EU average in terms of increase of demand of data workers in the period to 2025, in all scenarios. At the same time the U.K. education system increases its offering. The resulting data skills gap is mid-sized and lower than the average EU one as a share of demand in all scenarios.
- Both France and Germany show strong dynamics of growth of their national Data Markets as well as the number of data professionals in the observed period, resulting in high size gaps

[23] Eurostat has started collecting Tertiary Education statistics based on ISCED11. However, due to the many holes in the data series, this data has not yet been incorporated.

both in Baseline and High Growth scenarios. The German data skills gap is the largest in absolute size in the EU (up to 492,000 unfilled positions by 2025 in the High Growth scenario) but the French one is similar in size (485,000) and higher as a share of demand (25% in France vs 18% in Germany, High Growth scenario).

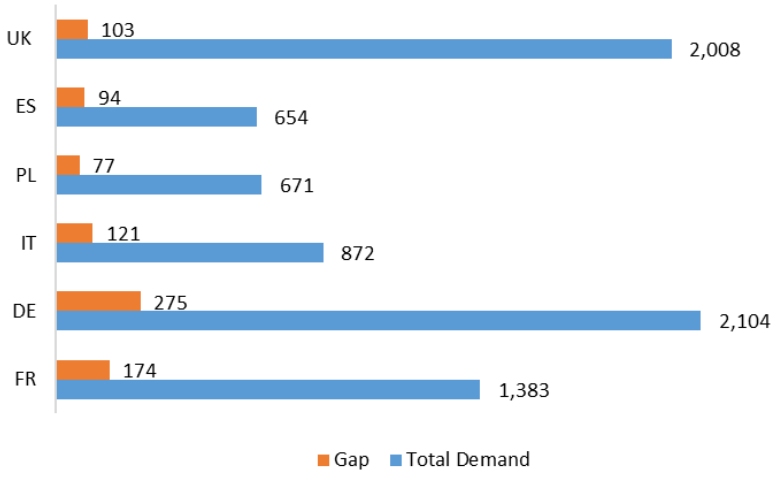
- Poland shows a data skills gap in line with those exhibited by most of the EU Member States across the three scenarios under consideration. The low growth of demand in the Challenge scenario results in a smaller gap than in 2016, while the opposite happens in the High growth scenario where the potential gap reaches a substantial level of total demand.
- Italy and Spain are characterised by dysfunctional labour markets, but their economies have started to grow again in the last 2 years and unemployment has gone down. Together with Poland they have a lower supply of graduates than the other large EU countries and rely more on inflows from other careers or upskilling, which is more elastic than education supply. This means that the variations of the balance demand-supply are higher by scenario and by year, with low data skills gaps in the Challenge scenario and high skills gap in the High Growth one.

*Table 51: Data Professionals Skills Gap by MS, 2016-2025 - Three Scenarios (000s)*

| | | FR | DE | IT | PL | ES | U.K. | Total Gap |
|---|---|---|---|---|---|---|---|---|
| 2016 | N. 000s | 40 | 54 | 53 | 75 | 35 | 85 | 428 |
| | % Gap | 4.9% | 4.2% | 9.3% | 14.7% | 8.65% | 6.0% | 6.2% |
| 2017 | N. 000s | 50 | 80 | 55 | 47 | 49 | 72 | 449 |
| | % Gap | 6% | 6% | 9% | 9% | 11% | 5% | 10% |
| 2020 | N. 000s | 84 | 156 | 82 | 54 | 55 | 97 | 699 |
| | % Gap | 8% | 10% | 12% | 9% | 11% | 6% | 8% |
| Baseline Scenario | 2025 N. 000s | 174 | 275 | 121 | 77 | 94 | 103 | 1,111 |
| | 17-25 CAGR | 17% | 16.7% | 10.4% | 6.4% | 8.6% | 4.7% | 12% |
| | % Gap | 11% | 12% | 12% | 10% | 13% | 5% | 10% |
| Challenge Scenario | 2025 N. 000s | 172 | 229 | 74 | 63 | 86 | 88 | 992 |
| | 17-25 CAGR | 16.8% | 14.1% | 3.7% | 3.9% | 7.4% | 2.6% | 10.4% |
| | % Gap | 11.3% | 9.9% | 8.7% | 9.2% | 12.5% | 4.8% | 9.4% |
| High Growth Scenario | 2025 N. 000s | 485 | 492 | 373 | 132 | 258 | 277 | 2,393 |
| | 17-25 CAGR | 32.9% | 25.5% | 27.1% | 13.9% | 23.1% | 18.5% | 23.3% |
| | % Gap | 25% | 18% | 29% | 15% | 27% | 11% | 18% |

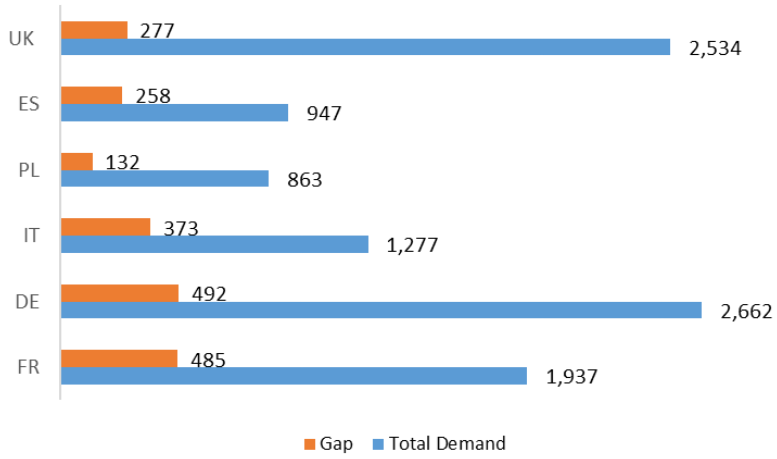
Source: European Data Market Monitoring Tool, IDC January 2018

*Figure 41: Data Skills Demand and Gap by MS, 2025, Baseline Scenario (000s)*



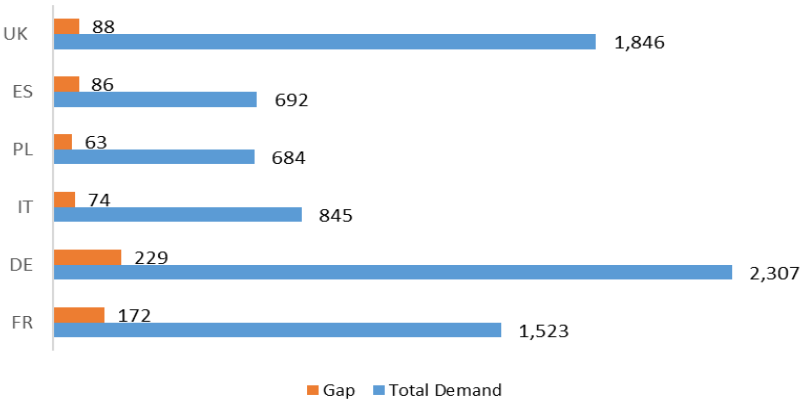

Source: European Data Market Monitoring Tool, IDC, January 2018

*Figure 42: Data Skills Demand and Gap by MS, 2025, High Growth Scenario (000s)*




Source: European Data Market Monitoring Tool, IDC, January 2018

*Figure 43: Data Skills Demand and Gap by MS, 2025, Challenge Scenario (000s)*




Source: European Data Market Monitoring Tool, IDC January 2018

## 9.5 Key Findings

According to this indicator, **there is a structural imbalance between demand and supply of data skills in Europe** since the first measurement of the EDM monitor for the year 2014.

**In the year 2017,** given the strong increase of demand of data professionals from 2016 (+7%), **the estimated gap grew by 10% reaching approximately 449,000 unfilled positions, corresponding to 6% of total demand**. By 2020 we expect the gap to expand to 699,000 unfilled positions in the EU27, corresponding to 8% of total demand.

**A vacancies ratio around 5% of demand or less is considered manageable.** From this point of view a data skills gap of 6% of demand as in 2017 is not a very large one, while the 8% gap foreseen for 2020 identifies stress in the market.

**The 3 scenarios highlight the diverse potential trajectories of the demand-supply balance of data skills to 2025**, predicting an increasing data skills gap, ranging from 10% of demand in the Challenge one, to 11% in the Baseline, to 19% in the High Growth (EU27). The slightly lower dynamic of data skills demand in the U.K., because of Brexit negative consequences, means that the gap incidence is smaller in the EU28.

**The absolute size of the data skills gap is also relevant, potentially reaching 1 million unfilled positions in 2025 in the EU27** Baseline scenario but up to over 2 million in the EU27 High Growth scenario. This raises difficulties for policies and initiatives aiming at addressing the gap.

**The gap indicator varies substantially by Member State in the forecast scenarios because the growth rates of both demand and supply can be very different**. Therefore, the gap indicator is very sensitive to specific national dynamics. For example, in Italy and Spain a lower elasticity of the higher education system results in potentially higher gaps in the High Growth scenario compared to the Baseline scenario.

**The gap forecasts extrapolate current supply and demand trends.** Actions to improve supply and make the market more efficient, reducing mismatches, can potentially reduce these gaps. The insufficient provision of data skills is a challenge particularly for the data industry which needs to hire highly specialised, technical data skills more difficult to find.

# 10. THE DATA ECONOMY BEYOND THE EU – US, BRAZIL AND JAPAN

While focusing primarily on Europe, the Data Market Monitoring Tool aims at providing quantitative evidence on the Data Market and the Data Economy in the European Union as a whole, on the individual Member States but also on some of the EU's main competitors worldwide. As a result, and in line with the outcomes produced in the course of the previous European Data Market Study (SMART 2013/0063), this report includes a specific chapter extending the analysis to the three main extra-European countries:

- The United States: the non-European country with the best innovation performance in 2017 according to the Global Innovation Index[24] as well as EU's top trading partner;
- Brazil: an upper-middle-income economy as well as one of the countries with the highest ICT Development Index (IDI) in Latin America;
- Japan: the largest high-income economy in the Asia/Pacific region, as well as the main Asian EU trading partner after China and a mature ICT market with a number of similarities with the EU as a whole.

Nevertheless, due to the wide disparity of the available statistical sources for these EU partners, we have kept our international focus on a restricted set of core indicators, in particular:

- Indicator 1.1: number of data professionals;
- Indicator 1.2: employment share of data professionals;
- Indicator 2.1 number of data companies;
- Indicator 3.1: revenues of data companies;
- Indicator 4: value of the Data Market;
- Indicator 5.1: value of the Data Economy;
- Indicator 5.2: incidence of the Data Economy on GDP.

## 10.1 Definition

The current estimate of the main indicators for the three select EU international partners leverages IDC databases available at worldwide level. Data such as ICT spending is available for most countries worldwide and is gathered with the same approach across the board – these data series are perfectly comparable at international level.

Other issues, however, continued to emerge when updating the indicators for the chosen EU international partners. For example:

- Unavailability of some specific data series used for the EU;
- Lack of information and data gathered with the survey, which was conducted only at European level.

IDC used existing data and desk research to estimate the key metrics for the three countries, accepting that categorisation of data professionals, data companies, and revenues of data companies might not be as current as for the EU Member States. We used IMF forecast data and

[24] (2017): The Global Innovation Index 2017: Innovation Feeding the World, Ithaca, Fontainebleau, and Geneva; https://www.globalinnovationindex.org/home

available statistics to validate our estimates. This is coupled with IDC's existing data sources of IT spending in the Black Book, and IDC's Spending Guides, both of which include spending for the three countries beyond the EU. The table below outlines the main sources used to estimate the international indicators in this report.

*Table 52: International Indicators: Main Data Sources*

| Data Source | Updated |
|---|---|
| IDC's worldwide and regional detailed market forecast estimates for IT Hardware, Software, and IT Services | Dec 2017 |
| IDC Worldwide Black Book (Standard Edition), quarterly updates from 2016 through December 2017. | Dec 2017 |
| IMF World Economic Outlook (WEO) Database | Oct 2017 |
| ILOSTAT statistics and databases | Jan 2018 |
| Consensus Forecasts, Consensus Economics, monthly updates, September 2016 – January 2018 | Jan 2018 |

## 10.2 The U.S.

The economy in the U.S. shows strong growth in 2017, and this is reflected in the increase the metrics for data professionals and the Data Economy. Organisations are investing in the Data Economy, using market strength in key innovation areas such as Artificial Intelligence and Cognitive Computing to offer additional services, extend the capabilities of many software tools that have interaction with customers or internal staff. Data streams from Internet of Things Devices, vehicles, mobile devices, and people interaction with companies push the capabilities of Big Data tools, and this encourages organisations to transform their businesses digitally, opening new markets and growing revenues. Larger organisations in the U.S. understand the value of the Data Economy and embrace it.

The positive development of the U.S.' Data Economy is confirmed by a solid year-on-year growth of the main indicators that we have tracked in this report, as well as in the previous European Data Market Study (SMART 2013/0063).

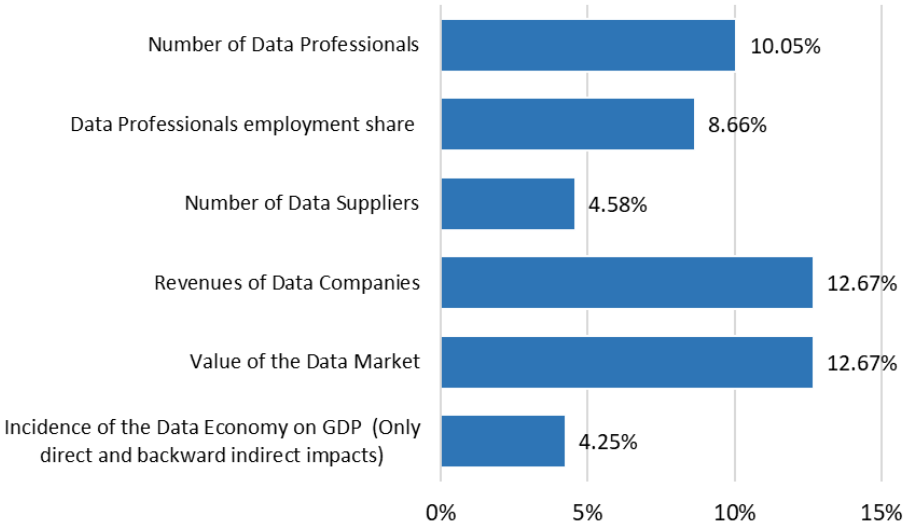
*Table 53: USA Indicators - Overview 2016, 2017*

| USA – Indicators' Overview | | | | | |
|---|---|---|---|---|---|
| N. | Name | Metrics | 2016 | 2017 | Growth 2017/2016 |
| 1.1 | Number of Data professionals | Total Number of Data professionals (Thousands) | 12,732 | 14,012 | 10.05% |
| 1.2 | Data professionals' employment share | % of Data professionals on total employment | 8.42% | 9.15% | 8.66% |
| 2.1 | Number of Data Suppliers | Total number of data supplier companies (000s) | 289,556 | 302,810 | 4.58% |
| 3.1 | Revenues of Data Companies | Total revenues generated by companies specialized in the supply of data-related products and services (Million €) | 129,173 | 145,546 | 12.67% |
| 4.1 | Value of the Data Market | Estimate of the overall a value of the Data Market (Million €) | 129,173 | 145,546 | 12.67% |
| 4.2 | Value of the Data Economy (Only Direct and Backward Indirect impacts) | Direct Impacts (Million €) | 108,521 | 113,677 | 4.75% |
| | | Backward Indirect Impacts (Million €) | 7,270 | 7,766 | 6.82% |
| 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% | 0.81% | 4.25% |

Source: European Data Market Monitoring Tool, IDC 2018

The number of data professionals, as well as their share on the country's total employment base, marks the strongest growth in 2017 vis-à-vis the other two international countries under consideration (10% and 8.7% year-on-year growth in the number of data professionals and employment share in 2017 over 2016 respectively). The same applies for the data supplier companies' indicators, with the highest increase of data suppliers in 2017 (4.6% vs. 1.1% and 3% in Brazil and Japan respectively) and an even stronger rise in their associated revenues and in the value of the Data Market in 2017 over 2016 (12.7% vs. 4.3% and 8.7% in Brazil and Japan respectively). Accordingly, the U.S.' overall Data Economy (direct and backward indirect impacts only) increased by almost 7% in 2017 gaining a 4.3% year-on-year in incidence on GDP and now representing more than 0.8% of the country's GDP. The figure below summarises this.

*Figure 44: USA Growth Rates 2017 - Key Metrics*




Source: European Data Market Monitoring Tool, IDC 2018

## 10.3 Brazil

Brazil experienced three years of weak growth, with 2015 and 2016 and part of 2017 marked by a significant GDP decline, according to the IMF. Despite the country's difficult macro-economic conjuncture and the current lack of investment in ICT and research & innovation,growth in data professionals, data companies, and the Data Economy has continued in Brazil over the past two years. This was and is inevitable because of the ongoing process of digital transformation at work globally and the benefits that come from digitally transforming companies. Brazil outlook is more positive for 2017, but the country still has some catching up to do when compared with European investment in the data economy.

*Table 54: Brazil Indicators - Overview 2016, 2017*

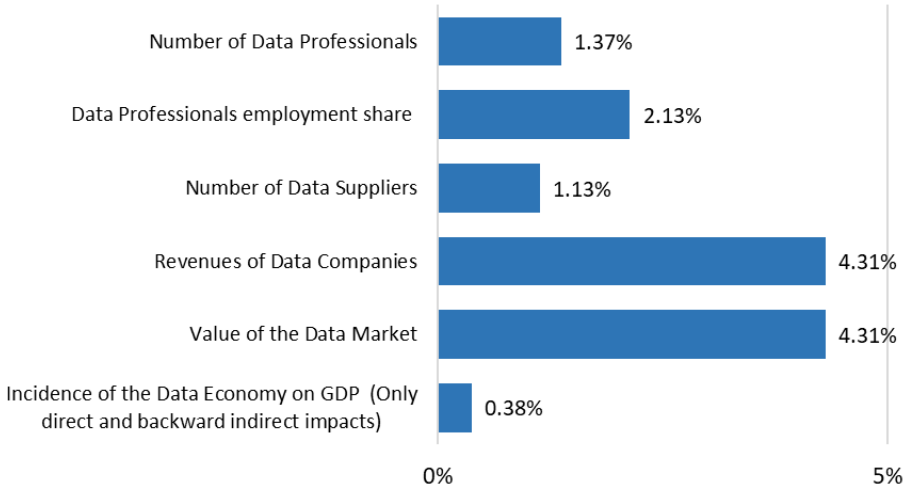
| Brazil – Indicators' Overview | | | | | |
|---|---|---|---|---|---|
| N. | Name | Metrics | 2016 | 2017 | Growth rate 2017/2016 |
| 1.1 | Number of Data professionals | Total Number of Data professionals (Thousands) | 1,160 | 1,176 | 1.37% |
| 1.2 | Data professionals' employment share | % of Data professionals on total employment | 1.81% | 1.84% | 2.13% |
| 2.1 | Number of Data Suppliers | Total number of data supplier companies (000s) | 35,979 | 36,387 | 1.13% |

| Brazil – Indicators' Overview | | | | | |
|---|---|---|---|---|---|
| 3.1 | Revenues of Data Companies | Total revenues generated by companies specialized in the supply of data-related products and services (Million €) | 6,049 | 6,310 | 4.31% |
| 4.1 | Value of the Data Market | Estimate of the overall a value of the Data Market (Million €) | 6,049 | 6,310 | 4.31% |
| 4.2 | Value of the Data Economy (Only Direct and Backward Indirect impacts) | Direct Impacts (Million €) | 6,157 | 6,395 | 3.86% |
| | | Backward Indirect Impacts (Million €) | 290 | 298 | 2.72% |
| 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% | 0.16% | 0.38% |

Source: European Data Market Monitoring Tool, IDC 2018

Gripped by corruption scandals, political instability and a severe loss of its economic dynamism, Brazil has had to come to terms with economic stagnation and a severe slowdown in its ICT investments. Nonetheless, the main indicators pertaining to the Data Market and the Data Economy have marked a growth in 2017 and are expected to continue on this journey in the years to come. Data professionals and their associated share on total employment have grown between 1.4% and 2.1%, the number of data supplier companies, their related revenues and the accompanying Data Market have grown of a considerable 4.3%. As a result, the incidence of the economy on Brazil's GDP will also increase but only by a modest 0.4% year-on-year (with the Data Economy – direct and backward indirect impacts - representing just 0.16% of Brazil's GDP in 2017). See the figure below for an overview of Brazil's growth rates in 2017.

*Figure 45: Brazil Growth Rates 2017 - Key Metrics*




Source: European Data Market Monitoring Tool, IDC 2018

## 10.4 Japan

Japan is on a renewed path of growth since the beginning of 2017 – ICT investments have resumed and the Data Economy as a whole has been consolidating its position and is now estimated at comparable levels with the European Data Economy. The number of data professionals continues to

rise at the same rate as for Europe, but the strength of other sectors of the economy in Japan means its share of GDP growth does not match the levels seen in the U.S. or in Europe.

Faced by both domestic and foreign critical issues (such as a slow inland demand, low internal consumption and difficult relationships with neighbouring North Korea and China), Japan has been able to recover and renew economic growth thanks, among other things, to a successful set of economic measures put in place by Prime Minister Shinzo Abe ("Abenomics").

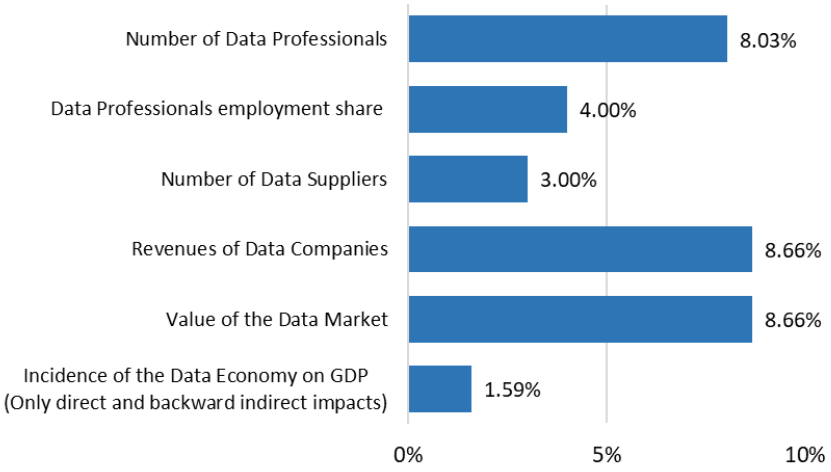
*Table 55: Japan Indicators - Overview 2016, 2017*

| Japan – Indicators' Overview | | | | | |
|---|---|---|---|---|---|
| N. | Name | Metrics | 2016 | 2017 | Growth rate 2017/2016 |
| 1.1 | Number of Data professionals | Total Number of Data professionals (Thousands) | 3,740 | 4,040 | 8.03% |
| 1.2 | Data professionals' employment share | % of Data professionals on total employment | 5.82% | 6.05% | 4.00% |
| 2.1 | Number of Data Suppliers | Total number of data supplier companies (000s) | 101,612 | 104,664 | 3.00% |
| 3.1 | Revenues of Data Companies | Total revenues generated by companies specialized in the supply of data-related products and services (Million €) | 25,513 | 27,723 | 8.66% |
| 4.1 | Value of the Data Market | Estimate of the overall a value of the Data Market (Million €) | 25,513 | 27,723 | 8.66% |
| 4.2 | Value of the Data Economy (Only Direct and Backward Indirect impacts) | Direct Impacts (Million €) | 27,394 | 29,949 | 9.33% |
| | | Backward Indirect Impacts (Million €) | 1,189 | 1,269 | 6.66% |
| 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% | 0.95% | 1.59% |

Source: European Data Market Monitoring Tool, IDC 2018

The indicators measuring the state of the Data Market and the Data Economy in Japan have all registered a significant growth in 2017 over the previous year. The number and employment share of data professionals has risen of 8% and 4% respectively in 2017, while the number of data suppliers has increased by 3% generating a significant growth in revenues and Data Market of 8.6%. The Data Economy has therefore also marked a positive development in Japan with an incidence on GDP now at 0.95% of the country's GDP – a growth of 1.6% with respect to the incidence in 2016. The figure below presents the main growth rates for Japan's Data Economy.

*Figure 46: Japan Growth Rates – Key Metrics*




Source: European Data Market Monitoring Tool, IDC 2018

## 10.5 International Overview and Comparison with the EU

The following tables and charts offer a comprehensive overview of some of the key indicators for the three international partners and the EU as a whole.

*Table 56: EU27 Indicators - Overview 2016, 2017*

Source: European Data Market Monitoring Tool, IDC 2018

| EU27 | | | | | |
|---|---|---|---|---|---|
| N. | Name | Metrics | 2016 | 2017 | Growth rate 2017/2016 |
| 1.1 | Number of Data Professionals | Total Number of Data Professionals (000s) | 4,875 | 5,273 | 8.20% |
| 2.1 | Number of Data Companies | Total number of data companies (000s) | 134,300 | 141,900 | 5.70% |
| 4.1 | Value of the Data Market | Estimate of the overall a value of the Data Market (Million €) | 46,183 | 50,438 | 12.67% |
| 4.2 | Value of the Data Economy (Only Direct and Backward Indirect impacts) | Direct Impacts (Million €) | 46,183 | 50,438 | 9.21% |
| | | Backward Indirect Direct Impacts (Million €) | 2,171 | 2,505 | 15.38% |
| 4.3 | Incidence of the Data Economy on GDP (Only direct and backward indirect impacts) | Ratio between value of the Data Economy and GDP (%) | 0.42% | 0.42% | 0% |

*Table 57: EU28 Indicators - Overview 2016, 2017*

| EU28 | | | | | |
|---|---|---|---|---|---|
| N. | Name | Metrics | 2016 | 2017 | Growth rate 2017/2016 |
| 1.1 | Number of Data Professionals | Total Number of Data Professionals (000s) | 6,891 | 7,290 | 8.00% |
| 2.1 | Number of Data Companies | Total number of data companies (000s) | 261,450 | 276,450 | 9.20% |
| 4.1 | Value of the Data Market | Estimate of the overall a value of the Data Market (Million €) | 59,496 | 65,038 | 9.30% |
| 4.2 | Value of the Data Economy (Only Direct and Backward Indirect impacts) | Direct Impacts (Million €) | 53,509 | 65,038 | 21.55% |
| | | Backward Indirect Impacts (Million €) | 2,780 | 3,303 | 18.79% |

| N. | Name | Metrics | EU28 2016 | EU28 2017 | EU28 Growth rate 2017/2016 |
|---|---|---|---|---|---|
| 4.3 | Incidence of the Data Economy on GDP (Only direct and backward indirect impacts) | Ratio between value of the Data Economy and GDP (%) | 0.42% | 0.52% | 23.81% |

Source: European Data Market Monitoring Tool, IDC 2018

*Table 58: Indicator 1.1 – Number of Data Professionals, International Overview and Comparison with the EU, 2016-2017 (000s)*

| Indicator 1.1 – Number of Data Professionals | | | |
|---|---|---|---|
| Country | 2016 | 2017 | Growth rate 2017/2016 |
| USA | 12,732 | 14,012 | 10.05% |
| Brazil | 1,160 | 1,176 | 1.37% |
| Japan | 3,740 | 4,040 | 8.03% |
| EU27 | 4,875 | 5,273 | 8.20% |
| EU28 | 6,891 | 7,290 | 8.00% |

Source: European Data Market Monitoring Tool, IDC 2018

*Table 59: Indicator 2.1 – Number of Data Companies, International Overview and Comparison with the EU, 2016-2017 (000s)*

| Indicator 2.1 – Number of Data Companies | | | |
|---|---|---|---|
| Country | 2016 | 2017 | Growth rate 2017/2016 |
| USA | 289,556 | 302,810 | 4.58% |
| Brazil | 35,979 | 36,387 | 1.13% |
| Japan | 101,612 | 104,664 | 3.00% |
| EU27 | 134,300 | 141,900 | 5.70% |
| EU28 | 261,450 | 276,450 | 9.20% |

Source: European Data Market Monitoring Tool, IDC 2018

*Table 60: Indicator 4.1 – Value of the Data Market, International Overview and Comparison with the EU, 2016-2017 (€, Million)*

| Indicator 4.1 – Value of the Data Market | | | |
|---|---|---|---|
| Country | 2016 | 2017 | Growth rate 2017/2016 |
| USA | 129,173 | 145,546 | 12.67% |
| Brazil | 6,049 | 6,310 | 4.31% |
| Japan | 25,513 | 27,723 | 8.66% |
| EU27 | 46,183 | 50,438 | 12.67% |
| EU28 | 59,496 | 65,038 | 9.30% |

Source: European Data Market Monitoring Tool, IDC 2018

*Table 61: Indicator 4.2 – Value of the Data Economy, International Overview and Comparison with the EU, 2016-2017 (€, Million)*

| Indicator 4.2 - Value of the Data Economy (Only Direct and Backward Indirect impacts) | | | | |
|---|---|---|---|---|
| Country | Metrics | 2016 | 2017 | Growth rate 2017/2016 |
| USA | Direct Impacts (Million €) | 108,521 | 113,677 | 4.75% |
| | Backward Indirect Impacts (Million €) | 7,270 | 7,766 | 6.82% |
| Brazil | Direct Impacts (Million €) | 6,157 | 6,395 | 3.86% |
| | Backward Indirect Impacts (Million €) | 290 | 298 | 2.72% |
| Japan | Direct Impacts (Million €) | 27,394 | 29,949 | 9.33% |
| | Backward Indirect Impacts (Million €) | 1,189 | 1,269 | 6.66% |
| EU27 | Direct Impacts (Million €) | 46,183 | 50,438 | 9.21% |
| | Backward Indirect Impacts (Million €) | 2,171 | 2,505 | 15.38% |
| EU28 | Direct Impacts (Million €) | 53,509 | 65,038 | 21.55% |
| | Backward Indirect Impacts (Million €) | 2,780 | 3,303 | 18.79% |

Source: European Data Market Monitoring Tool, IDC 2018

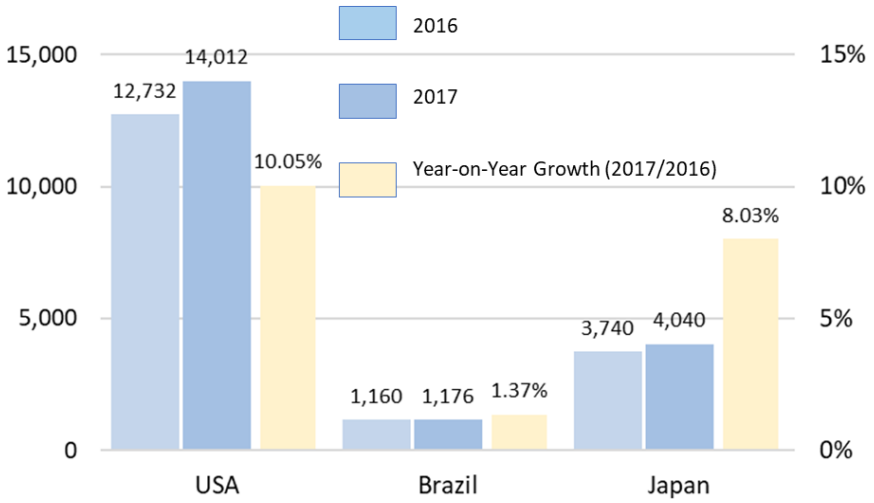
*Table 62: Indicator 4.3 – Incidence of the Data Economy on GDP (Only Direct and Backward Indirect Impacts), International Overview and Comparison with the EU, 2016-2017 (€, Million)*

| Indicator 4.3 – Incidence of the Data Economy on GDP | | | |
|---|---|---|---|
| Country | 2016 | 2017 | Growth rate 2017/2016 |
| USA | 0.78% | 0.81% | 4.25% |
| Brazil | 0.16% | 0.16% | 0.38% |
| Japan | 0.93% | 0.95% | 1.59% |
| EU27 | 0.42% | 0.42% | 0.00% |
| EU28 | 0.42% | 0.52% | 23.81% |

Source: European Data Market Monitoring Tool, IDC 2018

The U.S. confirm their leadership in the number and growth of data professionals – they are estimated at more than 14 million in 2017 with a year-on-year growth surpassing 10%. Brazil marks a very modest improvement due while Japan reaches the 4 million threshold and exhibits growth rates in line with those of Europe for both the EU27 and the EU28.

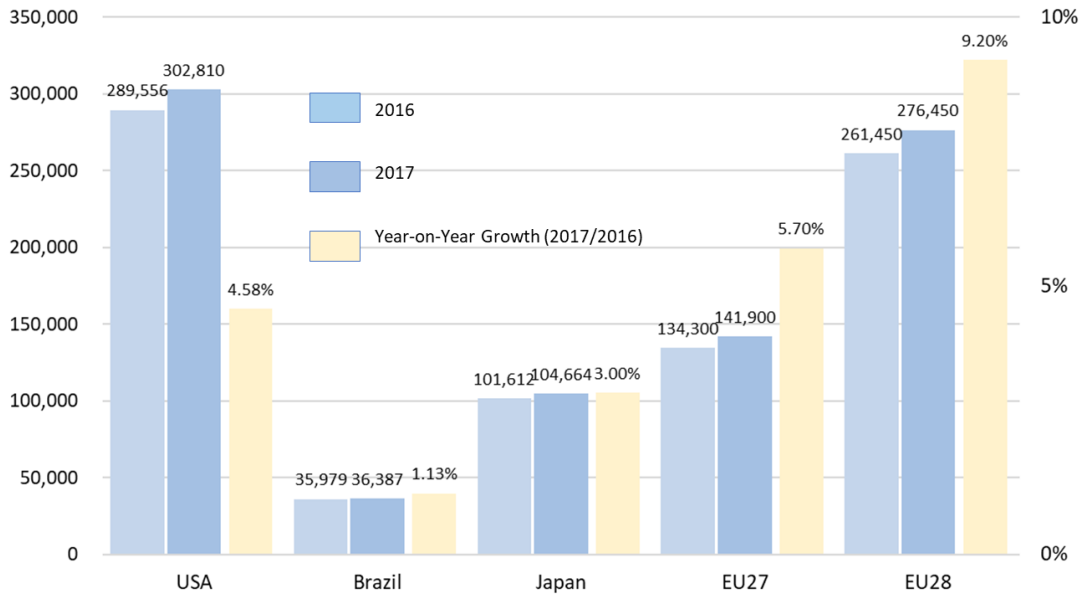
*Figure 47: Number of Data Professionals by Country, 2016-2017, Growth 2017 (Units; '000, %)*




Source: European Data Market Monitoring Tool, IDC 2018

Interestingly, though, Europe appears to be catching up on its gap with the U.S. and shows a renovated dynamism in some of the most significant data economy's areas. In terms of data suppliers (Figure 48) the EU can exhibit a year-on-year growth 2017-2016 of 9.2% - more than twice than in the U.S. and three times stronger than in Japan over the same period.

*Figure 48: Number of Data Suppliers in the U.S., Brazil, Japan and EU, 2016-2017, Growth 2017 (Units; '000, %)*

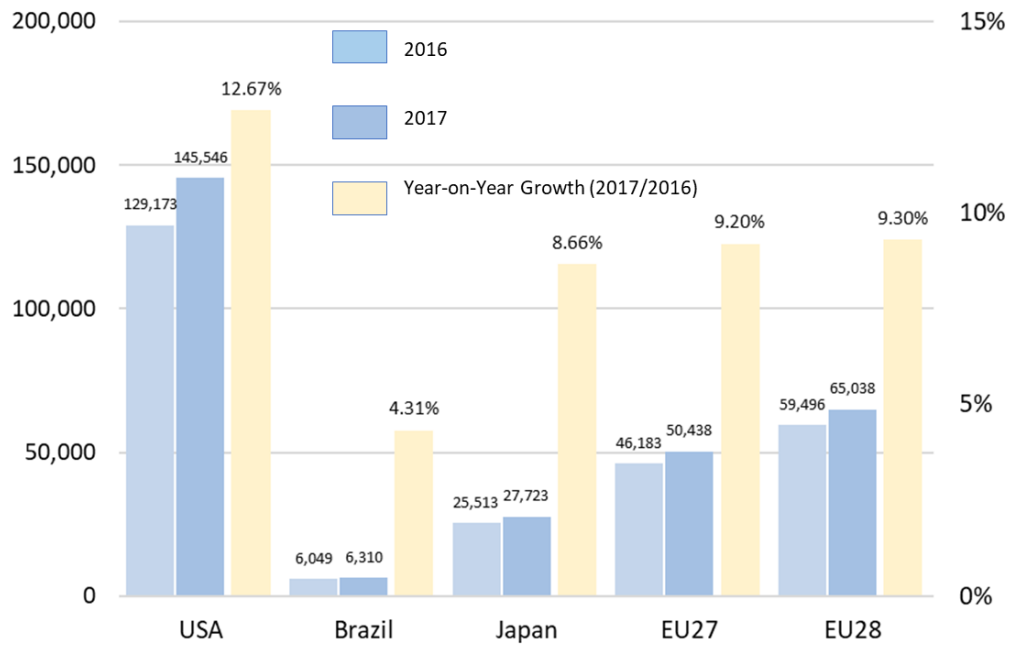



Source: European Data Market Monitoring Tool, IDC 2018

While close to double-digit growth, the value of the Data Market in the EU28 has marked a relative halt with respect to the U.S. in the period 2016-2017 but has continued to largely outpace Brazil and, to a much lesser extent, Japan. Indeed, the Data Market (Figure 49) will also see the U.S. in the

leading position with more than 145 million Euro in size and a buoyant year-on-year growth of 12.6% in 2017 over the previous year. In fact, the same dynamic applying to data professionals is at work for the market of data-driven products and services, with the EU being the only regional market in the position to challenge the American supremacy.

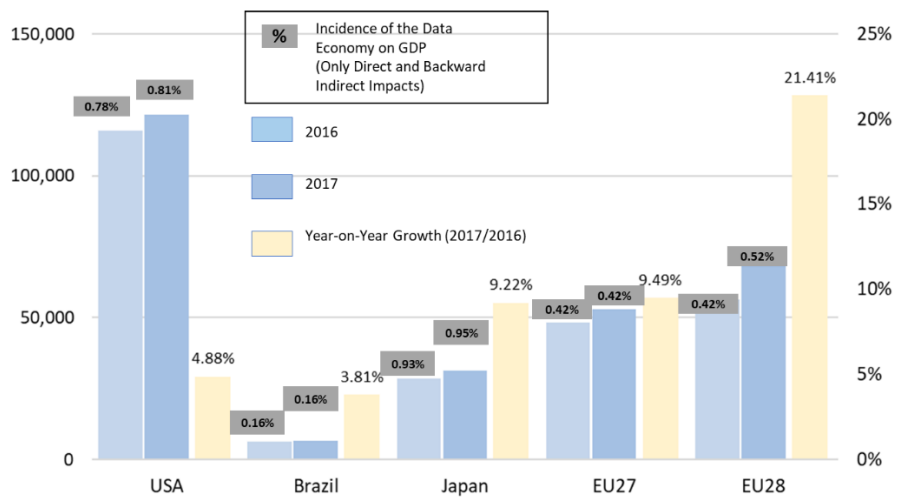
*Figure 49: Value and Growth of the Data Market in the U.S., Brazil, Japan and EU, 2016-2017, Growth 2017 (Units; '000, %)*




Source: European Data Market Monitoring Tool, IDC 2018

Figure 50 below displays the value of the Data Economy (direct and backward indirect impacts only) and its incidence on the GDP in 2016 and 2017. In line with the results obtained in the previous European Data Market Study (SMART 2013/0063), the U.S. continue to enjoy the highest impact of the Data Economy on their GDP - 0.81% in 2017, up 4.9% with respect to 2016. 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 50: Value, Growth and Incidence of the Data Economy as a Percentage of GDP in the U.S., Brazil, Japan and EU, 2016- 2017 (€, Million, %)*




Source: European Data Market Monitoring Tool, IDC 2018

## 10.6 Key Findings

**The U.S. confirm to have the healthiest Data Market and Data Economy** with consistent and sustained growth in the number of data professionals, companies, and the overall Data Market.

**The U.S. confirm their leadership in the number and growth of data professionals – they are estimated at more than 14 million in 2017 with a year-on-year growth surpassing 10%. Brazil marks a very modest improvement due while Japan reaches the 4 million threshold and exhibits growth rates in line with those of Europe for both the EU27 and the EU28.**

**Brazil shows lower growth following three difficult years for the country**, where investment in the Data Market took a lower priority. However, it is showing signs of recovery and has the potential for significant growth over the next few years.

**Japan's market is the closest to the European one in terms of growth and investment.** While regaining momentum, Japan's economy is still hampered by weak internal demand and lack of consumption. This, in turn, produces negative effects on ICT spending and, as a result, on limits the Data Economy and the Data Market potential.

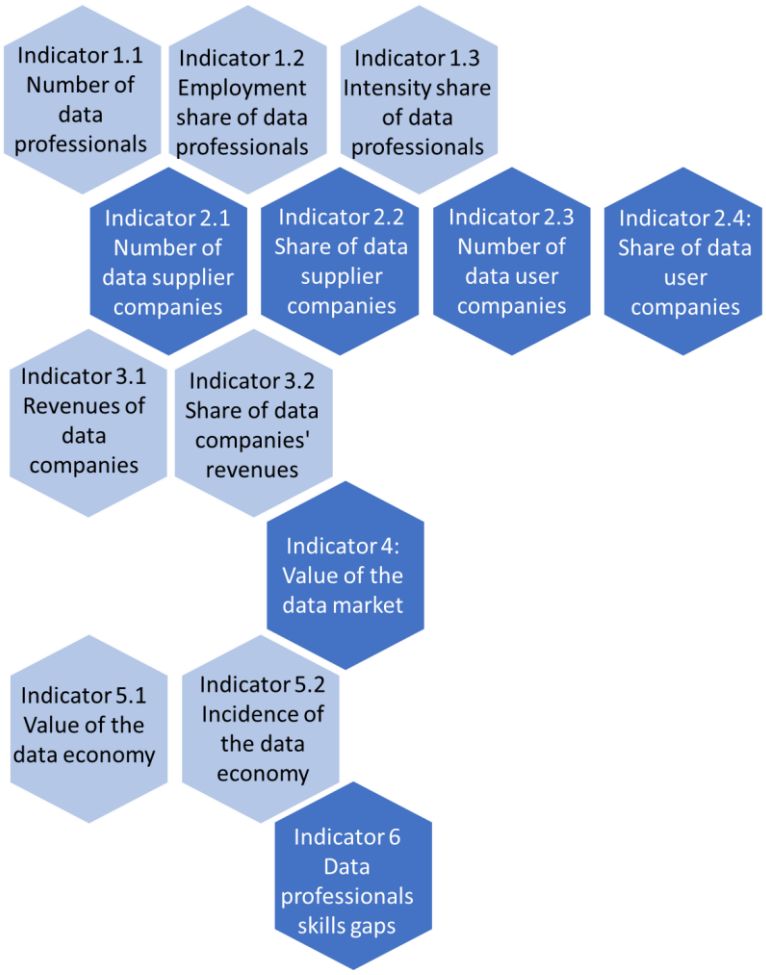
According to our first round of measurement of the international indicators, **the European Data Market and economy in the period 2016-2017 continues to consistently hold second place after the U.S. in value and, to a much lesser extent, in growth.**

**In terms of data suppliers, the EU exhibits a year-on-year growth 2017-2016 of 9.2% - more than twice than in the U.S. and three times stronger than in Japan over the same period.**

**Europe, however, 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.

# 11. CONCLUSIONS

This First Report on Facts & Figures (Deliverable D2.1 of the Update of the European Data Market Study, SMART 2016/0063) has presented the results obtained through the first round of measurements of the European Data Market Monitoring Tool for the period 2016-2017 with forecasts at 2025 under three distinct scenarios. The results pertained to the following set of indicators as per the updated version of the European Data Market Monitoring Tool:




Each indicator has been measured at the level of the total EU28 and for all 28 EU Member States, when available and applicable; industry-specific and company-size views are also offered with indicators provided by industry sector and company size bands, when possible. Two different views of the data are presented at European level: the usual EU28, and the EU27 excluding the United Kingdom. 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.

## 11.1 Data Professionals

The definition of the workers who collect, store, manage, analyse and visualise data as their primary or as a relevant part of their activity ("Data Professionals") has been substantially confirmed from the previous version of the European Data Market Study (SMART 2013/003) but partially renamed to better apprehend the current reality of the data-related workforce.

Data professionals are estimated at 5.3 million in the EU27 and 6.7 million in the EU28 in 2017, thus marking a significant increase in 2017 over the previous year (8.2% and 8% year-on-year respectively).

The employment share and the intensity share components of the data professionals' indicator are also expected to significantly improve in 2017 and 2020 if compared to our estimates in 2016 (now estimated at 3% and 3.4% in 20017 and 2020 in the EU27 and 3.2% and 3.6% for the same years in the EU28), thus confirming the positive evolution of the workforce involved in data-related professions over the period under consideration. This result shows direct continuity with the findings obtained in the previous European Data Market, SMART 2013/0063.

The average number of data professional per company will continue to grow steadily (although slowly) in 2017 and is estimated at an average of 11.5 per company by 2020 in the EU28.

The number of data professionals in the EU28 is expected to rise by a compound rate of 6.7% by 2025. This growth could be as high as 10.1% per year if the highest growth scenario is achieved – giving nearly 13.5 million data professionals in the EU28 and almost 11 million data professionals in the EU27 at 2025.

There is an even balance of data professionals' penetration across all Member States in 2017, aside from some unsurprising outliers which are data-intense economies (such as the U.K, Luxembourg or the Netherlands), or economies still struggling to make the most out of the ongoing digitization process (e.g.: Romania, Greece).

The highest concentration of data professionals is in the Professional services industry, followed by retail and wholesale. There is a notably wider spread of the number of data professionals by industry, although there is a fair correlation between the number of companies and employees in those industries and the number of data user companies in the same industries.

The highest growth in data professionals' year-on-year (in 2017 over 2016) is in the information and communications industry – although the growth is fairly even across all industries as each can benefit from transforming its business into a digital business.

## 11.2 Data Companies

**The organizations that are directly involved in the production, delivery and/or usage of data in the form of digital products, services and technologies can be both data suppliers and data users.**

The number of data suppliers continues to grow at a faster pace than the number of data users: the former is estimated at almost 142,000 in the EU27 and at more than 276,000 units in the EU28, thus exhibiting a year-on-year growth of 5.7% in 2017. Data users, instead, are projected to grow at 2.1% in 2017, amounting to almost 516,000 in the EU27 and to 690,650 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 more dynamic pictures of data companies in the EU, with growth rates constantly increasing over the past four years.

The number of data suppliers by Member State and by industry reflects the inherent business focus, and the strong financial, and services industries in the UK make the number of data suppliers in this Member State nearly five times that of the next biggest supplier of data tools and services.

By 2020 there will be nearly 722,000 data user companies – growing at a compound rate of 2.2%. This is faster than the growth of total companies. Which is expected to increase at less than 0.2% per year.

Data users' penetration rates (i.e. number of data user companies on total companies) vary significantly across Member States and present a wider range than the penetration rate exhibited by data professionals (i.e.: number of data professionals on total employment). With the an EU28 average standing at 6.6%, data user' penetration is as low as 1.0% for Romania, and as high as 12.2% for the Netherlands. This relates to the mix of industries associated with each country: those countries with a greater preponderance for data-oriented industries show a higher penetration and

will continue to do so. Adoption rates of data technologies are higher in industries such as Professional services, retail, and Financial services.

In Europe, there are approximately 33 accelerators and 14 incubators dealing with data-driven start-ups. According to the survey conducted in 2017 by the European Network of Centres of Excellence in Big Data Research, the European ecosystem comprises 251 data-driven start-ups and 611 SMEs. Based on evidence collected through desk-research and targeted interviews, there is a lively ecosystem in terms of supporting initiatives aimed at data-driven start-ups and innovative SMEs, including, but not limited to, business accelerators, incubators, venture capital funds and business angels' funds.

## 11.3 Data Companies' Revenues

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

Data companies' revenues will account for 3.2% of total company revenues in 2017, as companies continue their transition to the digital economy. As with other measures, there is a reasonable range of revenue shares across the Member States as those Member States with a stronger digital presence drive their economies' growth with digital services. Estonia, one of the leading adopters of digital technologies in business and commerce, will have 8.6% of its total company revenues accounted for by the Data Market, while Luxembourg has only 1.3%, and Ireland an even smaller 0.9%

In the different company size bands, it is the mid-sized companies that have the greatest revenues, as the mix of the number of companies and spending are maximized in this middle band. There are many more smaller companies, but as previously seen, these companies spend little on data tools and services, while the higher spend seen in the larger companies is offset by their low numbers – accounting for less than 1% of all companies. A clearer and more understandable picture is seen by looking at average company revenues by size band, where the larger companies show the greatest average spend.

Data 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 8.3% - easily outpacing the growth of the total ICT market over the same period (expected to be 1.6% from 2016 to 2025 baseline). The smaller Member States show the highest long-term growth as thy 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.

## 11.4 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 European Data Market will grow by 9.3% in 2017, and at an average rate of 6.0% out to 2020. This will make the market valued at close to 77.5 Euro Billion. Each of the Member States shows strong growth, well ahead of the expected growth for the IT market as a whole, which is projected to grow only 2.6% in 2017, and at an annual rate of 1.7% to 2020. The Data Market share of total ICT is 104% for 2017 and will grow out to 2025.

Estonia stands out as the Member State having a Data Market with the largest share of ICT spending; the country, however has embraced digital economics for some time, so its position is not surprising. Aside from Estonia, the spread of share of ICT spending taken by the Data Market is fairly narrow across the Member States with the largest share being 17 percent, and the smallest is 5.8%

The larger industries account for the largest share of the Data Market – as there is the greatest

number of companies within these industries, but adoption rates of Data Technology tend to be high in finance, Professional services, and in retail, which, together with the size of these industries makes them 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 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 2016 to 2025, with more than half coming from the three main industries of Manufacturing, Finance, and Professional services. The Data Market will continue to out-grow the total ICT market, with its share of this market rising from close to 10% in 2016, to more than 15% by 2025.

## 11.5 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.

Direct impacts will continue to grow and mirror the developments of the overall Data Market as measured by indicator 4. The indirect impacts will be the key impacts in all the scenarios, while induced impacts at the same time will start increasing, as the effects in the economy, through job and revenue creation as additional impacts will start to be visible.

In the baseline scenario the indirect and induced impacts will be more balanced than in 2017 and 2020. The overall impacts for EU27 will be 4.0% of the GDP. The indirect impacts will still play a fundamental role, as the user industries will consolidate the quantitative benefits from the use of data, but more than that, these benefits will go beyond the users and will translate in higher induced effects, generating jobs and revenues beyond the data companies itself.

With respect to the previous deliverable, from 2020 and 2025 there will be a change in the share of indirect impacts across the three scenarios. In particular, in the high growth scenario the penetration rate of data products and services into the user industry will be lower, and replaced by higher induced effects. The positive conditions under the high growth scenario will lead the overall impacts to exceed 1 trillion Euros in 2025.

Impacts by industry revealed that the financial sector, manufacturing, and professional services industries will all be heavily impacted by the exchange of data-driven products and services. Indeed, as also shown in the Data Market estimate, these industries make a significant usage of data-related technologies, their forward and backward impacts are high, and can convey effects at an induced level more than other industries thanks to several trends, such as the diffusion of IoT and Cloud computing, the digital transformation processes, as well as the usage of mobile and social technologies.

## 11.6 Data Professionals Skills Gap

The Data Professionals Skills Gap captures the potential gap between demand and supply of data skills in Europe. This is all the more relevant as the lack of skills may become a barrier to the development of Europe's data industry and the rapid adoption of data-driven innovation in the EU.

In the year 2017, given the strong increase of demand of data professionals from 2016 (+7%), the estimated gap grew by 10% reaching approximately 449,000 unfilled positions, corresponding to

6% of total demand. By 2020 we expect the gap to expand to 699,000 unfilled positions in the EU27, corresponding to 8% of total demand.

A vacancies ratio around 5% of demand or less is considered manageable. From this point of view a data skills gap of 6% of demand as in 2017 is not a very large one, while the 8% gap foreseen for 2020 identifies stress in the market.

The three scenarios highlight the diverse potential trajectories of the demand-supply balance of data skills to 2025, predicting an increasing data skills gap, ranging from 10% of demand in the Challenge one, to 11% in the Baseline, to 19% in the High Growth (EU27). The slightly lower dynamic of data skills demand in the UK, because of Brexit negative consequences, means that the gap incidence is smaller in the EU28.

The absolute size of the data skills gap is also relevant, potentially reaching 1 million unfilled positions in 2025 in the EU 27 Baseline scenario but up to over 2 million in the EU 27 High Growth scenario. This raises difficulties for policies and initiatives aiming at addressing the gap.

The gap forecasts extrapolate current supply and demand trends. Actions to improve supply and make the market more efficient, reducing mismatches, can potentially reduce these gaps. The insufficient provision of data skills is a challenge particularly for the data industry which needs to hire highly specialised, technical data skills more difficult to find.

## 11.7 Data Economy Beyond the EU – US, Brazil and Japan

This report extends the analysis of the Data Market and the Data Economy to three main extra-European countries: The United States, Brazil and Japan.

The U.S. confirm to have the healthiest Data Market and Data Economy with consistent and sustained growth in the number of data professionals, companies, and the overall Data Market.

Brazil shows lower growth following three difficult years for the country, where investment in the Data Market took a lower priority. However, it is showing signs of recovery and has the potential for significant growth over the next few years.

Japan's market is the closest to the European one in terms of growth and investment. While regaining momentum, Japan's economy is still hampered by weak internal demand and lack of consumption. This, in turn, produces negative effects on ICT spending and, as a result, on limits the Data Economy and the Data Market potential.

According to our first round of measurement of the international indicators, the European Data Market and economy in the period 2016-2017 continues to consistently hold second place after the U.S. in value and, to a much lesser extent, in growth.

Europe, however, 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.

# METHODOLOGICAL ANNEX

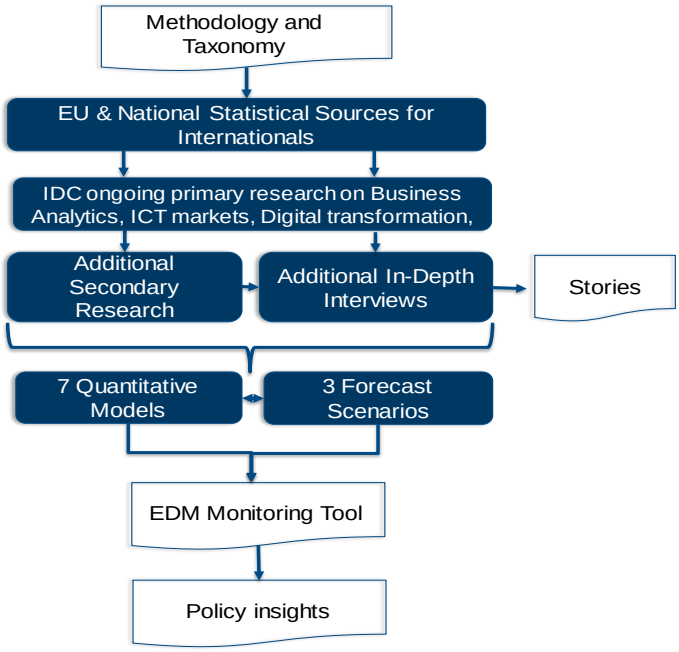
## Overview

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 51 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 51: 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 organized 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 will carry 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 organization 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 organization 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 organizations 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:

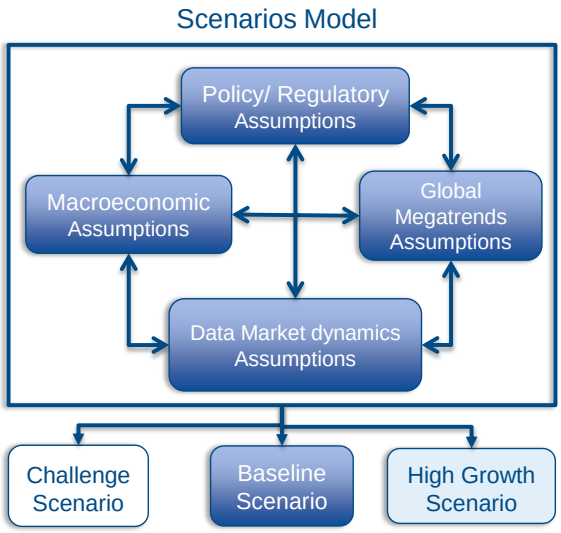
- High level of impact on the development of the Data Market;
- High level of uncertainty, with potential different outcomes (assumptions) over the next five years.

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 52). The scenarios are characterized 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 52: 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 63: 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:<br>EU GDP growth 2016-2025<br>ICT spending growth 2016-2025<br>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 64: 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 65: 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 66: Indicator 2 – Number of Data Companies*

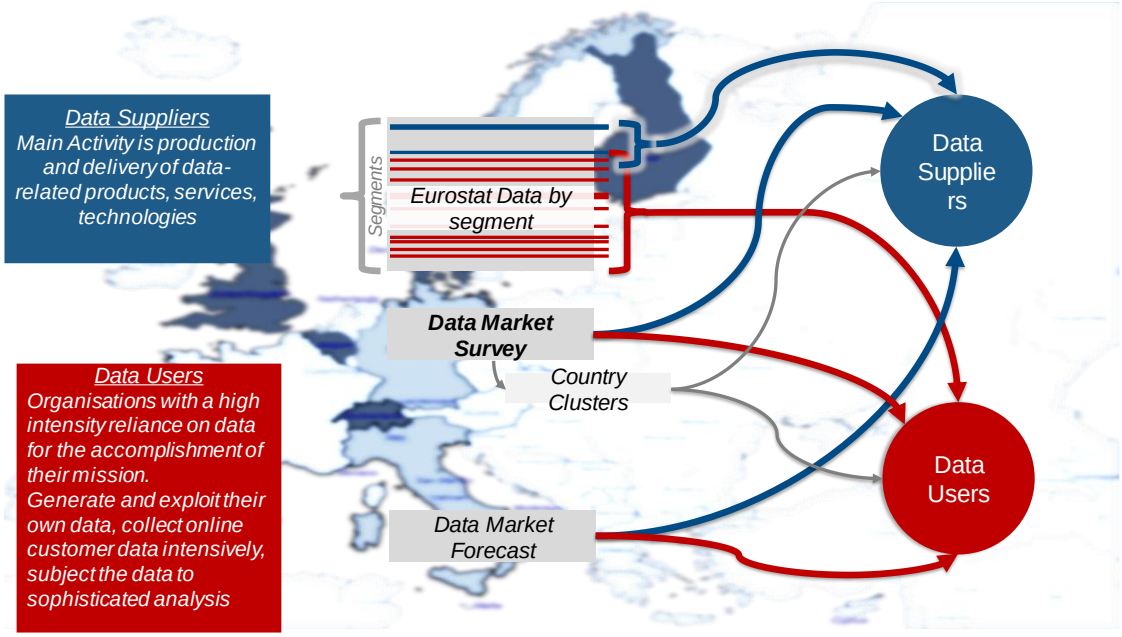
| Indicator 2 – Data companies | | | | |
|---|---|---|---|---|
| N. | Name | Description | Type and Time | Segmentation |
| 2.1 | Number of data suppliers | Total number of data suppliers, measured as legal entities based in the EU | Number, 2017-18-19<br>Forecast to 2025, 3 Scenarios | By Geography: 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:<br>below 250 employees<br>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<br>Forecast to 2025, 3 Scenarios | By Geography: 28 EU MS + total EU |
| | | | | By Industry: 11 industry sectors NACE rev.2 |
| | | | | By Company Size:<br>below 250 employees<br>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 53: 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 67: 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<br>Forecast to 2025, 3 Scenarios | By Geography: 28 EU MS + total EU<br>By Company Size:<br>below 250 employees<br>above 250 employees |
| 3.2 | Share of data companies' revenues | Total revenues of the Data Suppliers calculated by Indicator 2 | % of revenues on total,<br>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 68: Indicator 4 – Size of the Data Market*

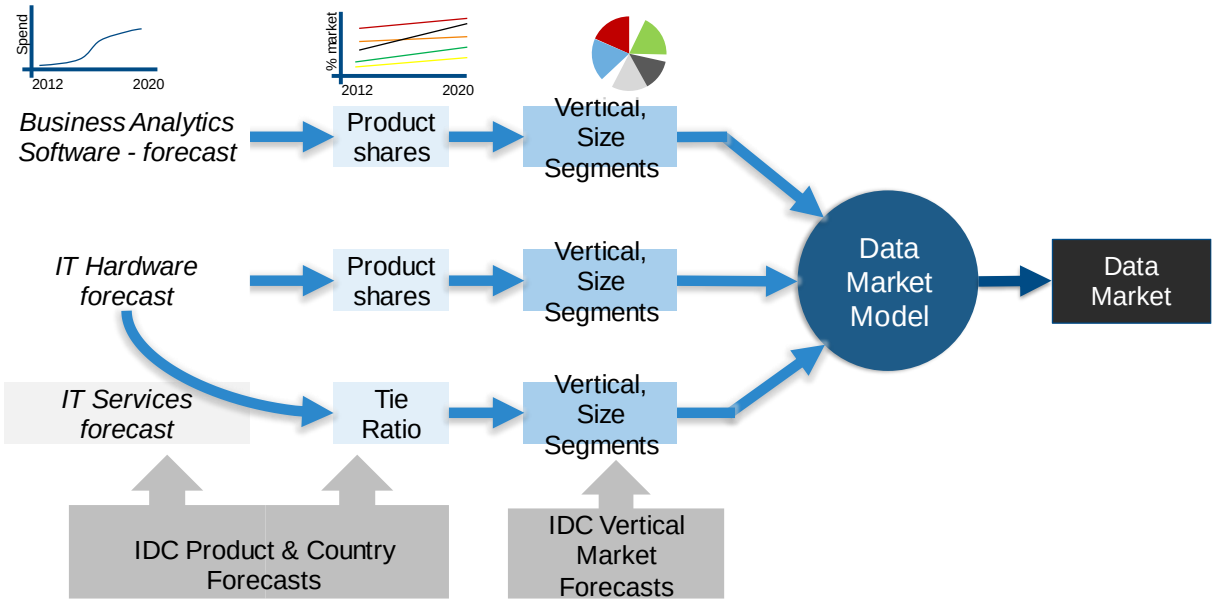
| Indicator 4 – Size of the Data Market | | | | |
|---|---|---|---|---|
| N. | Name | Description | Type and Time | Segmentation |
| 4 | Value of the Data Market | Estimate of the overall value of the Data Market | Billion €, 2017-18-19<br>Forecast to 2025, 3 Scenarios | By Geography: total EU, EU28<br>By Industry: 11 industry sectors NACE rev.2<br>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 54: 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 docoincide 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 69: 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<br>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<br>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 view 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 70: 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<br>Forecast to 2025, 3 scenarios | By Geography: total EU28; main EU Member States |

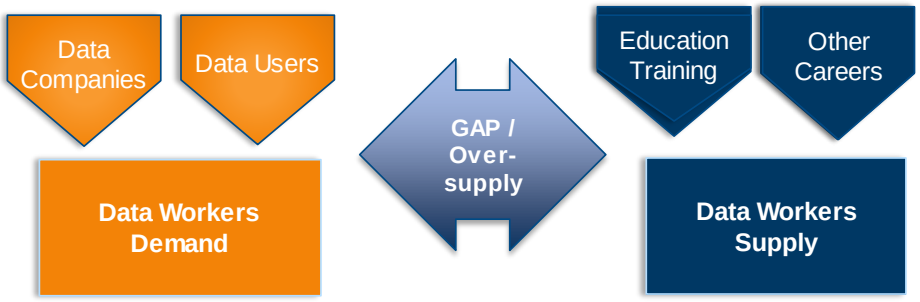
## *Methodology Approach*

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[25]. 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 55: The Data Skills Demand-Supply Balance Model*




Source: European Data Market Monitoring Tool, IDC 2016

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