European Data Market
SMART 2013/0063

D8 — Second Interim Report

The Data Market and The Data Economy

9th June 2016
<table>
<thead>
<tr>
<th><strong>Author(s)</strong></th>
<th>Gabriella Cattaneo, Mike Glennon, Rosanna Lifonti, Giorgio Micheletti, Alys Woodward, Marianne Kolding (IDC); David Osimo (Open Evidence)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Deliverable</strong></td>
<td>D8 — Second Interim Report (Second Release)</td>
</tr>
<tr>
<td><strong>Date of delivery</strong></td>
<td>9th June 2016</td>
</tr>
<tr>
<td><strong>Version</strong></td>
<td>3.0</td>
</tr>
</tbody>
</table>
| **Addressee officer** | Katalin IMREI  
Policy Officer  
European Commission, DG CONNECT  
Unit G3 — Data Value Chain  
EUFO 1/178, L-2557 Luxembourg/Gasperich  
katalin.imrei@ec.europa.eu |
| **Contract ref.** | N. 30-CE-0599839/00-39 |
1. Measuring the Data Market

1.1. 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.

In the present study, 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 (please see indicator 4.2 “Value of the Data Economy”). 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.

Table 28 provides an overview of the Data Market value in Europe in 2013, 2014 and 2015 and its forecast at 2020 according to the three scenarios under consideration.

Table 1 Indicator 4.1 Value and Growth of the Data Market

<table>
<thead>
<tr>
<th>N.</th>
<th>Name</th>
<th>Description</th>
<th>2013</th>
<th>2014*</th>
<th>2015</th>
<th>Growth rate 2015/2014</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Value of the Data Market</td>
<td>Estimate of the overall value of the Data Market</td>
<td>47,420</td>
<td>50,888</td>
<td>54,474</td>
<td>7.0%</td>
</tr>
</tbody>
</table>

Indicator 4.1 — Value and Growth of the Data Market (€ Million; %)

<table>
<thead>
<tr>
<th>N.</th>
<th>Name</th>
<th>Description</th>
<th>2020 Challenge (000)</th>
<th>2020 Baseline (000)</th>
<th>2020 High Growth (000)</th>
<th>CAGR 2020/2015 Challenge scenario</th>
<th>CAGR 2020/2015 Baseline scenario</th>
<th>CAGR 2020/2015 Growth scenario</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Value of the Data Market</td>
<td>Estimate of the overall value of the Data Market</td>
<td>72,086</td>
<td>84,001</td>
<td>111,924</td>
<td>5.8%</td>
<td>9.0%</td>
<td>15.5%</td>
</tr>
</tbody>
</table>

2014* = updated from the First Interim Report

Source: European Data Market Monitoring Tool, IDC 2016

The European data market in the EU28 is now estimated at EUR 50,888 million in 2014 - a slight increase of approximately 0.8% with respect to our previous estimate in the First Interim Report for the same year - and a buoyant increase of 7.3% with respect to the year 2013. In 2015, we expect the data market to reach EUR 54,474 million, marking a healthy growth rate of 7% year-on-year. The
growth will continue throughout the next five years, although at different paces according to the selected scenarios, registering a 9% growth rate under the Baseline scenario, a 15.5% under the High Growth scenario and a 5.8% under the Challenge scenario.

As a percentage of the total ICT spending in the EU28, the data market in 2014 is now estimated to represent a share of 8.7% (against the share of 8.6% that was calculated in the First Interim Report). This portion is projected to decrease somewhat in 2015 due to a considerable surge in the total ICT spending in that year. The data market share on the total ICT spending will however resume a solid growth in 2020 under all of the three scenarios considered in this study (10.6% under the Baseline scenario; 11.6% under the Challenge scenario and 12.1% in the High Growth scenario).

1.1.1. Updating the Measurement of the Data Market in 2015

IDC updated its forecasts for the data components used in estimating the size of the data market for this Second Interim Report. The update to IDC’s business Analytics forecast shows some small changes to forecasts for member states in the total out to 2018 – the limit for the previous forecast – and in some of the competitive markets used to model the Data market for the member states. IDC’s forecast for Spatial Information Analytics is lowered, but this constitutes less than three percent of the markets used to model the Data Market, Data Warehouse Generation Tools is lowered by ten percent by 2018, and Data Warehouse Management Tools are lowered by eleven percent. Austria has a lower expected share of those competitive markets contributing to the Data Market, as does Belgium, Finland, Romania and Sweden. However, in most cases these forecasts are lowered by 10-20 percent.

IDC raised its server hardware market forecast out to 2018 as a result of stronger markets in 2015 and out to 2018. Overall the server market is increased by 25 percent by 2018. Most Member States are impacted by this, but most affected are Austria, Denmark, Slovakia, and Slovenia. High End Enterprise Server shows the largest adjustment in 2015, rising by nearly 60 percent when compared with the previous forecast. However, this optimism wanes by 2018, with this revised forecast only 37 percent higher than the previous one used by 2018.

In contrast to this IDC’s revised storage forecast for the Member States shows little change out to 2018, increasing by less than one percent by 2018. A slightly higher forecast for entry level and high-end storage is tempered with a slightly lowered midrange storage forecast. These small changes effectively neutralise each other and the forecast is almost unchanged. Among the Member States, Denmark, Romania, and Slovakia have notably higher forecasts out to 2018, with IDC’s revised forecast close to 40 percent higher for these countries, while Finland and Greece have a lowered forecast for 2018 by 30-35 percent.

Each of these adjustments are explained in IDC’s published research and are included here to show their contribution to the revised Data Market forecast for this publication.

1.1.2. The Data Market and the Member States’ Dynamics in 2015

The data market value distribution by Member State in 2015 is exhibited in Figure 31. The overall picture does not change significantly from the one presented in our First Interim Report with the five largest EU economies (Germany, the UK, France, Italy and Spain) mustering almost 70% of the total data market value in the EU28.
In terms of year-on-year growth, Member States like the Netherlands, Sweden, France, the UK, the Czech Republic and Germany all exhibit a growth rate well above the EU28 average of 7.0% in 2015. In contrast, Italy, Portugal, Romania and Greece present a flat or negative growth year-on-year in 2014-2015.
Figure 2: Data Market Value (€M) by MS, 2015; Data Market Growth (%) by MS, 2015-2014

2014* = updated from the First Interim Report

Source: European Data Market Monitoring Tool, IDC 2016
The “catching-up” mode that we preconized in our First Interim Report is somehow becoming a reality: while still growing solidly, the smaller and ICT-advanced economies of northern Europe (e.g. Finland and Denmark) seem to have adopted a growth pace which is closer to the EU28 average, while larger but sometimes less agile economies like France, the UK and Germany are now reaping the benefits of their considerable size in term of ICT investments and the economy as a whole and have increased their data market spending above average in 2015.

Another way to look at these variables and further investigate the data market dynamic at Member State level is to compare the relative share of the data market share on the total EU28 with the relative ICT spending share as a percentage of the total EU ICT spending. As shown in Figure 26 below, the vast majority of the Member States presents a relative data market share that is still lower than their ICT spending share. In other words, the country’s relative weight at EU level in terms of data market is lesser than its relative weight in terms of ICT spending. In these Member States, the data market is expected to grow at fast pace in order to catch up with the ICT developments.

On the other hand, where the relative market share is higher than the relative ICT spending share, the data market has already achieved a significant level of development and it is likely to evolve steadily, although possibly at a slower speed, than in other Member States. This is notably the case of Germany and the UK.

*Figure 3 Data Market Share by MS (%) and ICT Spending Share by MS, 2015*

IDC research shows that the overall ICT spending in the EU in 2015 has considerably increased; as a result, the portion of the data market spending on the total ICT spending has slightly decreased representing 7.6% in 2015, down of one percentage point from the year 2014 as outlined in our First Interim Report (See Figure 27). In terms of country dynamics, the overall picture has not changed: Member States where the overall ICT spending is characterized by a solid component of Big Data and other data-related technologies and where investments in data tools are advanced display a penetration of data market spending on total ICT spending well above the EU28 average. In contrast, in those member states where ICT spending does not reach considerable amounts, the overall
expenditure for the data market – while sometimes considerable – remains in general around or below the EU average.

Figure 4 Data Market Share on ICT Spending by Member State, 2015, %

![Data Market Share Graph](image)

Source: European Data Market Monitoring Tool, IDC 2016

1.1.3. The Data Market and the Industry Dynamics in 2015

Figure 28 displays the value of the data market by industry and the relative growth by industry year-on-year across the EU28. The top industries in terms of data market size continue to be represented by industry sectors that make a significant usage of data-related technologies, i.e.: the ICT sector, the manufacturing sector, and the finance industry. In the previous year, our analysis found other segments at the top of the data market’s growing trend: Utilities and Healthcare (together with the ICT sector). Thus, unlike the past year, the most dynamic growth by industry in 2015 is exhibited by vertical markets that still hold a relatively smaller size of the overall data market spending by industry. This proves that the data-related technologies are rapidly finding new ground in previously unchartered areas and are growing fast in sectors like healthcare activities, professional services, home and consumers, and education.

Figure 5 Data Market Value by Industry, 2015, € Million; Data Market Growth by Industry, 2014-2015, %
As in 2014, the value of the data market by industry in 2015 is concentrated in four main sectors: the ICT sector, Finance, Manufacturing and the Public Sector (in particular, the Public Administration, defence and social security) muster two-thirds of the total data market value in 2015 (see Figure 29, below).

Figure 6 Data Market Value (€M) by Industry, 2015; Data Market Share (%) by Industry, 2015

2014* =

Source: European Data Market Monitoring Tool, IDC 2016
A somehow more interesting picture emerges when comparing the relative share of each industry in terms of the total data market spending with their respective relative share in terms of total ICT spending. The results for the year 2015 across the 12 industries under consideration are portrayed in Figure 30 below.

Similarly to what we experienced by Member State, the most interesting cases of data market dynamics by industry are where the relative share of data market spending is lower than the relative share of ICT spending as this is where potential growth is most likely to happen in the near future.

Conversely, where the relative share of data market spending is already higher than the relative share of ICT spending, data market growth tends to rely more on non-ICT factors, thus theoretically slowing down the growth pace of the data market as a whole.

Figure 7 Data Market Share by Industry (%) and ICT Spending Share by Industry, 2015
To further investigate the data market dynamics by industry, we present below our latest calculation of the share of the data market on the total of ICT spending in EU28 by industry. In 2015 the overall spending for the data market in the EU28 represented 7.7% of the total ICT spending in the EU28. By industry, this penetration rate takes the following form (see Figure 38).

Source: European Data Market Monitoring Tool, IDC 2016

Figure 8 Data Market Share on ICT Spending by Industry, 2015, %
Figure 31 above shows that, not surprisingly and again, the industries contributing the most to the overall ICT spending in Europe are often those displaying a higher data market spending with respect to the EU28 average in the EU. This is notably the case of the ICT sector, the transport sector, the education and the finance sectors. On the other end, industries like professional services, construction, and home, for example, lie well below the European average and constitute some of the most promising growth area in terms of data-related technologies over the next few years.

1.2. Indicator 4.1: Data Market Forecasts

The solid growth of the data market in 2020 as projected in our latest First Interim Report is confirmed in the current Second Interim Report. In fact, our estimates introduce slight improvements under all three scenarios with the Baseline scenario registering a 9% growth rate (against 8.7% estimated in the First Interim Report), the High Growth scenario displaying a 15.5% (14.1% in the First Interim Report) and the Challenge scenario still marking a considerable 5.8% growth (5.2% in the First Interim Report).

At Member State level our forecast for the period 2015-2020 under all three scenarios does not show any significant change to the picture we have drawn for 2014–2015. Figure 32 portrays the absolute size of the data market by Member State under the three scenarios. As outlined in the previous report, the size of the data market will continue to be closely interconnected with the Member States’ relative economic strength and the size of their ICT market. As a result, Germany, the U.K., France, Italy and Spain will continue to hold the lion’s share of the data market under the three scenarios.
A closer look at the Member States' growth dynamic reveals that the countries with a relatively high percentage of data market over their total ICT spending in the period 2015–2015 (see Figure 33) will continue to perform very well during the years 2015 to 2020.

Source: European Data Market Monitoring Tool, IDC 2015

Figure 10 Data Market Growth 2015-2020 by Member State, 2020, %, three scenarios
According to the Baseline scenario, highest growth continues to be expected in Member States characterized by substantial ICT spending growth, such as Sweden and Belgium. If compared to our previous estimates, smaller economies like Slovenia, Croatia and Slovakia will rapidly catch up with respect to their more advanced neighbours and reach over-the-average growth rates by 2020 in all three scenarios. This holds true under the positive High Growth scenario and the negative Challenge where Slovenia and Slovakia will exhibit powerful growth rates together with other Member States already performing very well according to our previous estimates (e.g.: Belgium, Sweden, the UK and Denmark). Greece will continue to struggle and suffer under all three scenarios and will be the only Member State featuring a negative growth rate under the Challenge scenario due to investment stagnation and the difficulty to overcome the overall economic hardship.

### 1.3. Key Findings

While still in its infancy, the European data market shows clear signs of consolidation and strengthening both at country and industry level.
• The European data market in the EU28 is now estimated at EUR 50,888 million in 2014 - a slight increase of approximately 0.8% with respect to our previous estimate in the First Interim Report for the same year - and a buoyant increase of 7.3% with respect to the year 2013 where it was estimated at EUR 47,420. In 2015, we expect the data market to reach EUR 54,474 million, marking a healthy growth rate of 7% year-on-year.

• As a percentage of the total ICT spending in the EU28, the data market in 2014 is now estimated to represent a share of 8.7% (against the share of 8.6% that was calculated in the First Interim Report). This portion is projected to decrease somewhat in 2015 due to a considerable surge in the total ICT spending in that year. The data market share on the total ICT spending will however resume a solid growth in 2020 under all of the three scenarios considered in this study (10.6% under the Baseline scenario; 11.6% under the Challenge scenario and 12.1% in the High Growth scenario).

• At Member State level, the market remains significantly concentrated within the largest economies of the EU with the UK, Germany, France, Italy and Spain making almost 70% of the overall data market in the EU28 in 2015. In terms of growth rate in 2015 over the previous year, Member States like the Netherlands, Sweden, France, the UK, the Czech Republic and Germany all exhibit a growth rate well above the EU28 average of 7.0% in 2015. In contrast, Italy, Portugal, Romania and Greece present a flat or negative growth year-on-year in 2014-2015. A possible interpretation of these results is that some of the largest (and sometimes less agile EU economies) – like France, the UK and Germany, are now starting to reap the benefits of their ICT investments in data-related technologies and have thus grown their data market spending well above average in 2015.

• The concentration rate of data market investments remains high with the ICT sector, Finance, Manufacturing and the Public Sector (in particular, the Public Administration, defence and social security) mustering almost two-thirds of the total data market value in 2015. The top industries in terms of data market size continue to be represented by industry sectors that make a significant usage of data-related technologies, although some changes are visible with respect to the previous year: while in 2014 for example utilities and healthcare year-on-year growth rate well above the EU average with respect to 2013, they have now resumed a growth pace much closer to the remaining industries. Unlike the past year, though, the most dynamic growth by industry in 2015 is exhibited by vertical markets that still hold a relatively smaller size of the overall data market spending by industry. This proves that the data-related technologies are rapidly finding new ground in previously unchartered areas and are growing fast in sectors like healthcare activities, professional services, home and consumers, and education.

• In terms of forecast at 2020, our current estimates do not diverge significantly from what we presented in our previous First Interim Report. As a result, the size of the data market will continue to be closely inter-connected with the Member States’ relative economic strength and the size of their ICT market. Consequently, Germany, the U.K., France, Italy and Spain will continue to hold the lion’s share of the data market under the three scenarios.

• According to the Baseline scenario, highest growth continues to be expected in Member States characterized by substantial ICT spending growth, such as Sweden and Belgium; smaller economies like Slovenia, Croatia and Slovakia will rapidly catch up with respect to their more advanced neighbours and reach over-the-average growth rates by 2020 in all three scenarios. This holds true under the positive High Growth scenario and the negative Challenge scenario. Greece will continue to struggle under all three scenarios and will be the only Member State featuring a negative growth rate under the Challenge scenario due to investment stagnation and the difficulty to overcome the overall economic hardship.
2. 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 also includes the direct, indirect, and induced effects of the data market on the economy. The definition of the data economy remains therefore unaltered and is conformant to the one in use in our First Interim Report.

Indicator 4.2 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 will then be 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.

2. **The indirect impacts**: indirect impacts are all the impacts which take place in other industries related to the considered industry, in our case the data industry. 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 depending on the adoption of data by the downstream industries, i.e. the data users as classified in a selected number of industries as explained in the Methodology Report (Annex 3). 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 created by additional payment of wages to staff in the data industry and its direct supply chain. A proportion of this will be spent on consumer goods and services. This leads to further business growth throughout the EU economy. The IDC model quantifies the economic impacts on other industries as data workers spend their earnings. The additional consumption of data workers and of data companies' suppliers will in fact support economic activity in various industries such as retail, consumer goods, banks, entertainment, etc.
As in the previous report, our 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.

2.1.1. Updating the Value of the Data Economy for the year 2015

The benefits related to the use of data products and services have been explored in our survey conducted between the end of the year 2014 and the beginning of 2015. In the survey, we asked about the benefits typology with reference to the year 2014 and about the benefits in quantitative terms for the year 2014 and 2015. We have assumed that there were substantially no changes in the kind of benefits the companies were having from the year 2014 to the year 2015. Besides, the survey asked about quantitative benefits for the year 2014, while for the year 2015 we asked what was expected in terms of quantitative benefits. For the estimates of the data products and services impacts we used both the data provided by the survey: the impacts registered by companies in 2014 and the benefits expected for the year 2015. We have assumed that the expectations of the companies were correct and matched the actual benefits. We believe by the way that this is a reasonable assumption since the market did not change substantially or in an unexpected way from 2014 to 2015.

To finalize our estimates of the impacts in 2015, the study team has also conducted additional desk research on a series of different internal and external sources, among which:

- 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
- IMF World Economic Outlook (WEO) Database, April 2016

8.2 Indicator 4.2 and 4.3: Value of the Data Economy and Incidence of the Data Economy on GDP

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2</td>
<td>Value of the Data Economy</td>
<td>Value of the Data Market plus estimate of direct, indirect and induced impacts on the economy</td>
<td>246,840</td>
<td>257,589</td>
<td>272,047</td>
<td>4.4%</td>
<td>5.6%</td>
<td>1.83%</td>
<td>1.85%</td>
<td>1.87%</td>
</tr>
</tbody>
</table>

2014* = updated from the First Interim Report
**Indicator 4.2 — Value and Impact on GDP of the Data Economy in 2020, Three Scenarios (€ Million; %)**

<table>
<thead>
<tr>
<th>N.</th>
<th>Name</th>
<th>Description</th>
<th>2020 Challenge (000)</th>
<th>2020 Baseline (000)</th>
<th>2020 High Growth (000)</th>
<th>Impact on GDP 2020 Challenge</th>
<th>Impact on GDP 2020 Baseline</th>
<th>Impact on GDP 2020 High-Growth</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Value of the Data Economy</td>
<td>Value of the Data Market plus estimate of direct, indirect and induced impacts on the economy</td>
<td>405,109</td>
<td>643,115</td>
<td>1,075,988</td>
<td>2.64%</td>
<td>3.17%</td>
<td>5.14%</td>
</tr>
</tbody>
</table>

Source: European Data Market Monitoring Tool, IDC 2016

The overall value of the data economy was worth almost EUR 247 Billion in 2013 in the EU and remains at a level of more than EUR 257 Billion in 2014, representing over 1.85% of the European GDP and is estimated at 272 Billion in 2015 representing 1.87% of GDP and marking a growth rate year-on-year of 5.6%.

There are 13 countries with an impact of the data economy over the EU average. Among these countries we find, not surprisingly, Germany, the UK, Finland as well as countries like Estonia and Finland.

Major impacts can be found in Member States where:

- Data products and services have a high penetration in the domestic economy so that the benefits from using data and introducing innovation are widespread.

- Companies adopting data products and services have greater benefits than companies in other countries.

What is more, the introduction and diffusion of new technologies - like data technologies – is likely to provide immediate and sizeable impacts to those economies that are relatively new adopters of innovation: this is notably the case of Malta, Cyprus, and Lithuania.

Total indirect impacts are the most important component of the data economy as they involve backward and forward industries. The forward indirect impacts are vital as they involve all data user industries, which, as explained in the Methodology Report, are a significant part of the overall industry. The indirect impacts represent in fact 55% of the total impacts, and in some countries nearly around 60% - a portion that remains nearly constant over the three years under consideration. The impacts of the diffusion are not producing major changes in terms of impacts. However, we should not conclude that the diffusion of data products and services is not having a significant impact in the short term: two-three years is a too short period to measure the full potential and impacts for an emergent technology.
The indirect impacts are strongly correlated with the data market: where the demand is more developed the impacts produced through users are more relevant in terms of absolute value. We should also notice that the total impacts on GDP varies from 1% to about 2.5%. There are specifically some small countries with relevant impacts such as Malta, Estonia, Latvia, Hungary, Finland, while large countries in terms of GDP and of data market show a total impact on GDP around 1%-1.5%. This may depend on the fact that in some small countries the adoption of data products and services is faster than it is in other countries so that the users involved have relevant benefits.

Furthermore, we should be aware that the growth rate for the accumulated impacts may not be as fast as the growth rate for the data market because of the following reasons:

- The diffusion of innovation evens the competitive advantages of the adopters, and prices of the new technologies and products tend to decrease. This has to be considered an advantage for consumers, also in terms of quality of processes and products, but it does not necessarily translate into an increase in revenues and GDP.

- Data should be considered to be a multipurpose innovation, which in part may be a substitute for other products and technologies so that the net effects may decrease over time.

Also, the overall impacts of the data economy very much depend on penetration rates of data products and services. Since these rates are nearly stable over the observed years, so are the impacts relatively stable. What really changes the incidence of impacts is a widespread diffusion of data-driven innovation throughout the whole the economic system. As explained, we are currently in the first stages of the technology diffusion curve with penetration rates logically following a rather moderate trend.
The overall induced impacts (over EUR 63 Billion in 2013, EUR 66 Billion in 2014 and almost EUR 73 Billion in 2015) represent approximately 0.5% of European GDP; we have adopted a conservative approach for the estimate of the induced impacts due to two main reasons:

- Data workers do not necessarily correspond to new jobs: therefore, all the data workers are not necessarily going to earn additional wages;
- During an economic crisis, the propensity to save increases, which may extend the typical time-lag of such impacts (they tend to occur later than the market launch of new products and services).

### Figure 12 Impacts of the Data Economy by MS as a % of total impacts, 2015

![Figure 12: Impacts of the Data Economy by MS as a % of total impacts, 2015](image)

Source: European Data Market Monitoring Tool, IDC 2016

Overall, the indirect impacts, including forwards and backwards, are the most relevant impacts, although this is not true in all countries. The countries where the indirect impacts are more relevant are the one leading the data market.

The induced impacts are always around 0.5% of the GDP. This estimate also depends on the long economic crisis the EU is experiencing, as explained above and on the fact that the data workers, until now, are not necessarily corresponding to new jobs.

As a whole, the total impacts of the data products and services are relevant in terms of GDP and correspond to the benefits deployed by the most innovative products.

### 8.3 Analysis and Interpretation

The amount of data generated, collected, managed and elaborated through data analytics tools is exponentially increasing. Such an amount of data and its management are dramatically changing the knowledge process creation of social and economic systems and the decision making processes as well. Analysts underlined that the new decision-making processes act as a rationalization and optimization factor (Brynjolfsson, 2011, Mc Kinsey, 2012), since they improve effectiveness and efficiency, and in some cases they may have a disruptive effect. The impacts related to the new decision-making processes are the one we have called the forward indirect impacts.
The value creation process based on data rests on the elaboration of information and knowledge (OECD 2016), although the boundaries between data, information, and knowledge are sometimes fuzzy. The huge volume of data is a global phenomenon which is sometimes viewed with suspicion by citizens, consumers, and businesses because data flows are seen as an intrusion of the privacy. Nevertheless, there is currently some evidence showing that data analysis can provide benefits to both businesses and consumers. By the way, this is not surprising since we should remind that the economic theory holds that information encourages competition between businesses for the benefit of consumers.

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

- **Extracting information from structured and unstructured data**: data analytics techniques are today able to analyse both structured and unstructured data. We should remind here that most data stored by businesses are unstructured (IDC, 2012). Technologies such as optical character recognition, natural language processing, face recognition algorithms, and machine learning algorithms are empowering the use of all data.

- **Real-time monitoring and tracking**: analysis of data in real time is often mentioned as one of the most powerful factors 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 specialized data analytic 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 healthcare 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. This is by the way out of our field of analysis, but since we are dealing with data impacts we are going to present the most relevant ways in which data provide benefits to organisations (McKinsey, 2012, OECD, xxx). 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 kinds 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 allow 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. We should stress here that these are also the more significant impacts in terms of GDP.

*Figure 13 Data economy: Distribution of impacts as a % of total impacts, 2015*

![Data economy: Distribution of impacts as a % of total impacts, 2015](image)

Source: European Data Market Monitoring Tool, IDC 2016
8.4 Indicator 4.2: Data Economy Forecasts

The survey carried out for this study looked only at forecasts for the impacts of the data economy on the overall EU economy and on the "big six" European Member States.

In the First Interim Report we estimated the impacts of the data economy only for the “big six”. In this Second Interim Report we estimate the impacts of the data economy for all the 28 EU Member States. As stated in the assumptions, we assumed that a number of countries are and will have similar impacts as, in average, the main countries. For the forecasts, we also assumed that:

- the impacts are positively correlated with the increase of the market in general
- the impacts are going to be strongly related with the penetration rate of the data products and services within the economic system.

The total indirect impacts (including backward and forward industries) of the data industry itself, are the most relevant effects, followed by the induced impacts.

To analyse the impacts at 2020, we estimated the impacts within three different scenarios: Challenge, Baseline and High Growth. The scenarios are fully explained in Chapter 3 of the present report. The scenarios are based on different macro-economic trends. 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.

Figure 14 Impacts of the Data Economy by MS as a % of total impacts, 2020 Challenge Scenario

![Graph showing impacts by MS as a % of total impacts](image)

Source: European Data Market Monitoring Tool, IDC 2016

In the Baseline scenario, on the opposite, Europe will recover from the crisis showing a GDP trend 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.
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 demand data products and services more than they did into the Baseline scenario 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, and the penetration rate in the user industry will be twice the penetration rate of the Baseline scenario.

The High Growth scenario, therefore, explores the impacts of data products and services with a GDP equal to the one in the Baseline scenario, but with an increased penetration of data services and products.

The total impact of data products and services at 2020 may vary, in the three scenarios, from 2.6% of the GDP in the Challenge scenario to nearly 3.2% in the Baseline one and to 5.1% in the High Growth scenario.
The forward indirect impacts are confirmed to be the most important. These impacts are all the impacts deployed through the user industries. In 2020, and within the Baseline scenario, the forward indirect impacts will be supported by an increasing penetration rate of data as well as by increasing benefits due to gains in efficiency and competitiveness in the user companies.

We have confirmed a conservative estimate of the induced impact because, as explained in the data worker chapter, data workers do not correspond to net job creation and a share of the data workers are people already working in the companies, dealing with management or ICT.

In the Challenge scenario, the forecast value of the data economy will represent about 2.6% of GDP of the EU GDP. This narrow difference between the Baseline scenario (3.2% of GDP) and the Challenge scenario depends on the fact that the penetration rate between the two scenarios is narrow. This means that the data diffusion is at a stage where the diffusion of such new products and services will occur in both scenarios. We need to note here that our Challenge scenario is a moderately pessimistic one which is not going to stop the innovation diffusion especially where innovation is able to provide benefits in terms of more efficiency and competitiveness.

Besides, the High Growth scenario is characterised by a much higher penetration of data into the user industry. The High Growth scenario shows an impact on GDP of 5.1%. This relevant impact is mainly driven by the forward indirect impact as it was in the other two scenarios. Nevertheless, in this case, the forward indirect impacts represent 67% of the total impact, more than it was into the other scenarios. This effect can easily be explained: in this scenario, the penetration rate of data services and products into the user industry is much faster than it was into the other scenarios; this is the reason why in the High Growth scenario the impacts achieved through the users are more relevant than they were in the other scenarios. The multiplier effect of data product and services is higher when, with equal macroeconomic conditions, the penetration in the user industry is higher.

On the overall, we confirm the forecasts already presented last year, although the current ones are a bit more positive, due to the faster increase forecasted for the data market.

When looking at the impacts at country level, we observe that in the baseline scenario, there are 10 countries with a forecasted impact on GDP higher than the average forecasted for the EU, while the other 18 countries register a lower incidence of the impacts on GDP. The countries with the higher share of impacts on GDP are the countries where the demand for data products and services will grow with a fast trend. Where the penetration rate is faster, the impacts on GDP will be even more relevant.

We should here remind that the market can increase because of a higher intensity of demand (the same companies demand more products) or because of a faster penetration (more companies demand more data products). Based on our analysis, it seems that, at the same level of the total demand, a higher penetration rate will produce higher impacts, measured as a % of the GDP.

### 8.5 Key Findings

- The economic impact analysis is an effective tool for the scrutiny of the impacts of a multipurpose and widespread innovation such as the data products and services.

This kind of approach helps subdivide the impacts in order to better understand the source of such impacts, and whether they can be considered new additional impacts on the economic system. An in-depth and reliable analysis focusing on impacts should be based on specific field research, but the survey provided interesting insights about the impacts gathered by the companies adopting data products and services. A significant share of the companies is starting to see benefits in terms of additional revenues.
• In 2013 the total impacts estimated for data products and services represented EUR 246,840 Million, which is equivalent to 1.83% of the EU GDP; in 2014 EUR 257,589 Million, which is equivalent to 1.85% of EU GDP. In 2015, we estimate the same indicator to be at EUR 272,047 Million, representing 1.87%

The qualitative analysis of the impacts, based on the survey results (specifically for the years 2013 and 2014), as well as the IDC cases and the stories collected for this study, shows that the main impacts in terms of GDP are the forward indirect impacts, which are the impacts deployed on the user industries. The type of innovation we are dealing with and its adoption process also suggest that this impact is the "actual" new impact emerging in the economic system.

• The impacts gathered by the user industries (forward indirect impacts) represented almost 56% of the total impacts in 2013, approximately 54% in 2014 and remain substantially unchanged in 2015, corresponding to 1% of total EU GDP.

When compared with the total impacts in 2013 (1.83% of GDP), we can see that the total impact is not increasing very fast and significantly year on year. This is quite normal and in line with the overall impacts of innovations. First of all, because the penetration of the data products and services is not so fast and also because the growth rate for the accumulated impacts may not be as fast as the growth rate for the data market (see 8.1).

The user industries are starting to see the quantitative benefits from the use of data. These benefits on the user industry translate into revenue increases. In turn these lead to GDP increases in the region of 1% for the EU.

• The scenarios at 2020 show that a high penetration in the user industry produces relevant and fast impacts in terms of GDP.

The High Growth scenario, under similar macroeconomic conditions, produces relevant impacts on the user industries. A fast penetration of data products and services produces relevant effects in terms of GDP though the benefits achieved by the user industry. This means that policies that leverage on increasing demand for products and services may provide relevant impacts.