EUROPEAN DATA PORTAL

Re-using Open Data

A study on companies transforming Open Data into economic & societal value

This study has been prepared by Capgemini Invent as part of the European Data Portal. The European Data Portal is an initiative of the European Commission, implemented with the support of a consortium[i] led by Capgemini Invent, including Intrasoft International, Fraunhofer Fokus, con.terra, Sogeti, 52North, Time.Lex, the Lisbon Council, and the University of Southampton. The Publications Office of the European Union is responsible for contract management of the European Data Portal.

For more information about this paper, please contact:

European Commission
Directorate General for Communications Networks, Content and Technology
Unit G.1 Data Policy and Innovation
Daniele Rizzi – Policy Officer
Email: daniele.rizzi@ec.europa.eu

European Data Portal
Gianfranco Cecconi, European Data Portal Lead
Email: gianfranco.cecconi@capgemini.com

Written by:
Jorn Berends
Wendy Carrara
Wander Engbers
Heleen Vollers

Last update: 15.07.2020
www: https://europeandataportal.eu/
@: info@europeandataportal.eu

DISCLAIMER
By the European Commission, Directorate-General of Communications Networks, Content and Technology. The information and views set out in this publication are those of the author(s) and do not necessarily reflect the official opinion of the Commission. The Commission does not guarantee the accuracy of the data included in this study. Neither the Commission nor any person acting on the Commission's behalf may be held responsible for the use, which may be made of the information contained therein.

Luxembourg: Publications Office of the European Union, 2020

OA-03-20-042-EN-N ISBN: 978-92-78-41872-4 doi: 10.2830/876679

[i] At the time this report was first issued the consortium consisted of: Capgemini Invent, Intrasoft International, Fraunhofer Fokus, con.terra, Sogeti, the Open Data Institute, Time.Lex, and the University of Southampton.

Abstract

Well-established businesses use Open Data to enhance their existing services, to optimise their processes or to provide insights derived from data analysis to help solve clients' business challenges. 23% of the organisations expect their turnover related to Open Data to grow between 11% and 60% in the coming years, while 37% of the respondents forecast their turnover to increase by at least 61% annually.

The objective of this report is to understand how organisations, primarily within the private sector, use Open Data and what business models have been developed around its re-use. One of the key findings is the mismatch between available data sets and the data sets that are most re-used. In order to access Open Data, national Open Data portals are used most, followed by directly accessing the public administrations themselves.

External as well as internal barriers do remain and hinder re-users to standardise or automate the collection and processing of Open Data. The report concludes on a series of recommendations for both the public and private sector. The public sector needs to better align their data provision strategy with the specific needs of users. The private sector is invited to share further success stories around the re-use of Open Data and demonstrate how it benefits growth and innovation.

Résumé

Les entreprises établies depuis quelques temps vont principalement faire usage des données ouvertes dans le but d'améliorer leurs services existants, leur process ou encore travailler à la résolution des défis posés par leurs clients. Dans les 5 années à venir, 23% des organisations interrogées s'attendent à une augmentation de 11 à 60% de la part de leur chiffre d'affaire lié à l'utilisation de données ouvertes. Cette croissance est estimée à plus de 61% par plus d'un tiers des entreprises interrogées.

Cette étude vise à comprendre comment les différentes organisations utilisent les données ouvertes et quel modèle économique elles ont développé dans ce sens, et ce plus particulièrement dans le secteur privé. L'étude a permis de révéler l'inadéquation entre l'offre de données et les jeux de données les plus utilisés. L'accès aux données ouvertes se pratique désormais davantage par le biais de portails nationaux, ou bien encore auprès des administrations publiques.

Toutes fois, des freins internes et externes limitent encore l'utilisation des données, notamment dans le domaine de la standardisation et de la collecte automatique de données. Le rapport conclut sur une liste de recommandations s'adressant à la fois aux utilisateurs de données mais aussi aux administrations publiques qui les fournissent. Le secteur public est appelé à mieux aligner l'offre de données avec la demande exprimée par le secteur privé et associatif. Le secteur privé, quant à lui est invité à partager davantage de belles histoires sur l'utilisation des données et les bénéfices liés à la croissance et à l'innovation.

Table of Contents

Table of Contents 3

List of figures 5

Executive summary 7

1. Introduction 14

2. Why Open Data Matters 17

2.1 The Open Data Value Chain 17

2.2 The Economic Impact of Open Data 18

2.2.1 Value created through Open Data in the EU28+ 18

2.2.2 Societal benefits of Open Data 21

2.3 Understanding the benefits of Open Data at company level 22

3. What data do businesses re-use? 25

3.1 Which data categories are the most re-used? 25

3.2 Where does the Data come from? 33

3.3 The access to Open Data 35

3.4 The benefits of working with Open Data 37

4. What business is there with Open Data? 41

4.1 What business does Open Data bring? 41

4.1.1 Selling services as the most important source of revenue 41

4.1.2 Who are the clients? 45

4.1.3 Large discrepancies remain when investigating share of Open Data in turnover 46

4.2 The role of Open Data in business models 49

4.2.1 The emergence of data brokers 49

4.2.2 Organisations are combining business models 51

4.2.3 Adding value with the use of Open Data 53

4.2.4 Web shops and websites are the most important distribution mechanisms 58

4.3 The Open Data capacity of organisations 59

5. What skills are needed to work with Open Data? 61

5.1 Profiles companies are seeking to hire 61

5.2 How to acquire these skills? 65

6. The way forward 66

6.1 The main barriers encountered by businesses 66

6.2 Success factors for working with Open Data 71

6.3 What do organisations using Open Data want for the future? 76

7. Conclusion 78
Annex I: Methods used 79
1. Online survey 79
2. Interviews 80
Selection process 80
Methodology interview 79
3. IODC Temperature Check 79
Annex II: List of organisations showcased 81
Annex III: Correlation Matrix data categories 86
Annex IV: Interview questionnaire 87
Annex V: Survey questionnaire 89
Annex VI: Temperature check during IODC 100

List of figures

Figure 1: Economic value of Open Data 7
Figure 2: Three step approach to the study 8
Figure 3: Sources of revenue for the organisations working with Open Data (in percentages) 9
Figure 4: Survey respondents according to the Open Data Value Chain Archetypes (in percentages) 10
Figure 5: Profiles of recruits Open Data organisations are looking for (in percentages) 10
Figure 6: Critical elements for organisations (in percentages). 11
Figure 7: Open Data Maturity Clusters 14
Figure 8: Structure of the report 15
Figure 9: Open Data Value Chain 17
Figure 10: Data Value Chain Archetypes 18
Figure 11: Direct market size of Open Data per country 19
Figure 12: Forecasted share of direct Open Data in EU 28+ GDP, 2016-2020 19
Figure 13: Total market size of Open Data per market sector for EU28+ in millions, 2020 20
Figure 14: Forecasted Total number of direct Open Data jobs (in person), 2016- 2020 20
Figure 15: Examples of non-economic benefits of Open Data 21
Figure 16: Three Pillar Approach 22
Figure 17: Country distribution of assessed respondents (in absolute numbers) 23
Figure 18: Type of organisation – IODC Temperature Check (in percentages) 23
Figure 19: Most used data categories by Open Data organisations (in percentages) 26
Figure 20: Re-used data categories vs. most available data sets by data category on the European Data Portal (in percentages) 27
Figure 21: High priority data domains 28
Figure 22: Data domains re-used most by respondents (in percentages) 28
Figure 23: Most consulted data domains by the EU28+ countries (in absolute numbers). High priority data domains appear in orange 29
Figure 24: Correlation matrix of the re-use of Open Data categories 30
Figure 25: Main data categories used – IODC Temperature Check 32
Figure 26: Countries from which Open Data is used (in percentages) 33
Figure 27: Regions Open Data is used from (in percentages) 34
Figure 28: Regions from which Open Data is used (in absolute numbers) 35
Figure 29: Platforms via which data is accessed (in percentages) 36
Figure 30: Main benefits of working with Open Data (in percentages) 38
Figure 31: Main Benefits - Public Sector - Private Sector - NGOs/Not for profit – IODC Temperature Check (in percentages) 38
Figure 32: Sources of revenue for the Open Data organisations (in percentages) 41
Figure 33: Clustering of qualitative descriptions of revenue streams 42
Figure 34: Business activities of Open Data organisations (in percentages) 44
Figure 35: Client segments 45
Figure 36: Turnover details of Open Data organisations (in percentages) 46
Figure 37: Percentage of turnover related to Open Data (in percentages) 47
Figure 38: Forecasted annual growth in general and related to Open Data (in percentages) 48
Figure 39: Forecasted general growth compared to the growth related to Open Data 49
Figure 40: Respondents as archetypes of Open Data 50

Figure 41: Archetypes - Private sector - Public Sector - NGOs/Not for for profit – IODC Temperature Check (in percentages) 50
Figure 42: Mentioned business models (in percentages) 52
Figure 43: Organisations using one or multiple business models (in percentages) 52
Figure 44: Comparison between archetypes and business models (in percentages) 53
Figure 45: Distribution mechanisms (in percentages) 58
Figure 46: Number of internal and external employees with a full time equivalent workload (in absolute numbers) 59
Figure 47: Companies foreseeing to recruit (in percentages) 60
Figure 48: Profiles respondents are looking for (in percentage) 61
Figure 49: required e-competences for a data scientist 62
Figure 50: Skills sought for in employees (in percentages) 63
Figure 51: Skills required to work with Open Data 64
Figure 52: Correlation matrix of skills (correlations > 0.1 are marked in orange, correlations < -0.1 in blue) 64
Figure 53: Skills sought for most often (in percentages) 65
Figure 54: Main Challenges - Public Sector - Private Sector - NGOs/Not for profit IODC Temperature Check (in percentages) 69
Figure 55: Critical elements for organisations (in percentages) 72
Figure 56: Language preference when re-using data sets (in percentages) 73
Figure 57: Does language matter? (in percentages) 74
Figure 58: Company ownership (in percentages) 79
Figure 59: Founding year of all valid entries (in absolute numbers) 80
Figure 60: Geographical and regional spread of the EU respondents (in absolute numbers) 81
Figure 61: Overview of the NACE Codes per country (in absolute numbers) 78
Figure 62: Interviewed organisations 79
Figure 63: Type of organisations – IODC Temperature Check 80
Figure 64: Correlation matrix data categories 86

Executive summary

The potential value of Open Data can only be unlocked if the data is made available in the first place. European countries have been developing a number of policies as well as data portals that are increasingly sophisticated. The launch of the European Data Portal in November 2015 improves access to data, whatever its country of origin and whatever its language.

The objective of the present study is to understand how organisations, primarily within the private sector, use Open Data and what business models have been developed around its re-use. With better insight into how Open Data is transformed into value, two ambitions of the European Data Portal can be strengthened:

- Helping Open Data providers with a better understanding of how to use Open Data which can help them adapt their data provision strategy and thereby improve the (re-)usability of public sector information.
- Broadening the knowledgebase demonstrating the potential of the Open Data for society at large, thanks to a collection of stories from organisations that use Open Data to create value and, finally, show the variety of what can be achieved by using Open Data.

The first step in assessing the microeconomic benefits of Open Data, is to provide an overview of the macroeconomic benefits of Open Data. The market size of Open Data is expected to increase by 36.9% from 2016 to 2020 to a value of 75.7 bn EUR in 2020[1]. The forecasted number of direct Open Data jobs is expected to rise from 75,000 in 2016 to nearly 100,000 jobs by 2020. Thanks to the positive economic effect on innovation and the development of numerous tools to increase efficiency, not only the private sector, but also the public sector is expected to experience an increased level of cost savings through the re-use of Open Data to a total of 1.7 bn EUR in the EU28+ by 2020. The figure below summarises the key figures per economic indicator.

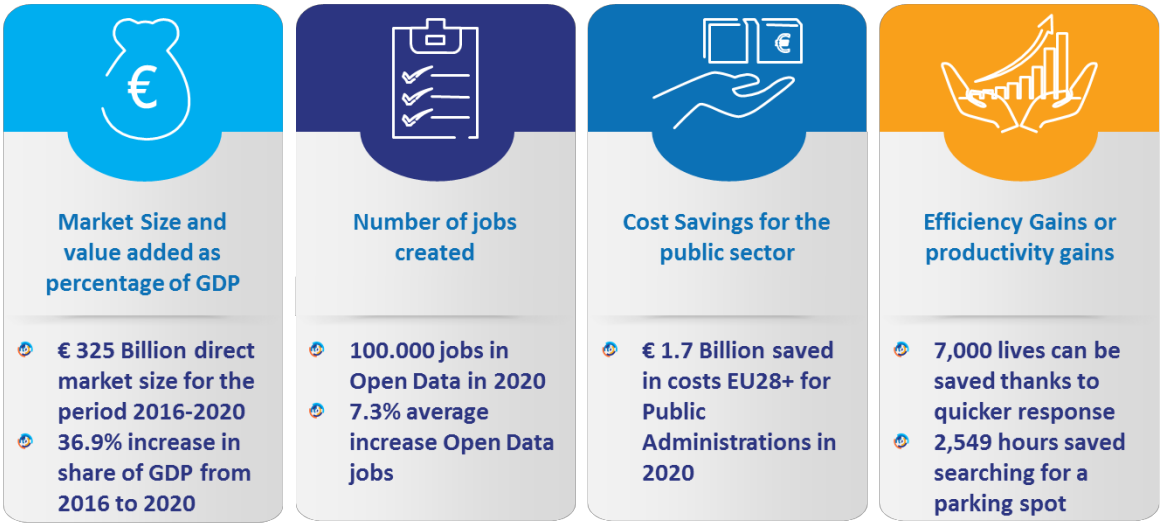


Figure 1: Economic value of Open Data

In order to further explore the use of Open Data at company level, a survey was conducted inviting organisations from across the Europe to share their story on how they use Open Data to create value. The stories collected via a survey are used to shed light on how companies transform Open Data from raw material into a service or product. To gain more insight in the particular challenges and success

[1] European Data Portal, 2015, Creating Value through Open Data

stories of organisations working with open data, in-depth interviews were conducted with a selection of organisations. In order to complement and cross-check these findings, a final survey was conducted at the International Open Data Conference that took place in October 2016. In addition, the feedback mechanism available on the European Data Portal allows companies to share their story on how they re-use Open Data[2].

Figure 2: Three step approach to the study

The organisations that responded to the survey are based in 21 countries of which six are non-European. Spain and the United Kingdom are overrepresented in the sample and correspond to countries that are considered mature in the field of Open Data[3]. With 85% of the respondents, the majority of the organisations was established in the 2000s of which 28% in 2014. The organisations are relatively small, with over 60% of them having 0-5 employees. This does not mean that larger corporations are ignoring Open Data. It may underline the fact that they may not wish to openly communicate about how they use Open Data. Following the survey, in-depth interviews are conducted with 33 organisations, with a solid spread across European countries and NACE-sectors. In a final stage, the respondents of the Temperature Check (see section 2.3) belong to the public sector (32%), the private sector (36%) and the non-government/not-for-profit sector (32%).

The use and application of Open Data is characterised by a wide diversity, while the organisations themselves show similarities. Most of the organisations that use Open Data as a resource are relatively young and can be identified as start-ups in which leadership often conjuncts with the ownership of the company. Most of them are active in the information and communication sector. One of the key findings when looking into what data these organisations re-use is the mismatch between available data sets and the data sets that are most re-used. Public administrations publish much data about the environment and the legal system and & public safety, yet most data re-users are more interested in data about the government & public sector, economy & finance, transport and more particularly business registries and company data. National Open Data Portals are the most used platforms to access Open Data. However, despite the growth in data portals, the second most used source to collect data are still public administrations themselves. This could indicate a preference for the Open Data users to get their data from a platform close to the source where the data is produced.

While most organisations serve clients from multiple sectors, their own economic activity is predominantly found in data processing or analysis. The principal source of revenue for most organisations is the provision of services to their clients, followed by selling products. Services provided are for instance consulting services to transform raw data into actionable insights or the provision of Software as a Service (SaaS) to process and analyse data.

[2] European Data Portal, 2016, Tell us your story

[3] European Data Portal, 2016, Open Data Maturity in Europe

Figure 3 shows that the majority of the organisations working with Open Data obtains their revenue by selling services or products or a combination of both. The category 'Other' includes other ways of generating revenue, such as selling products & advertisements or donations received by NGOs.

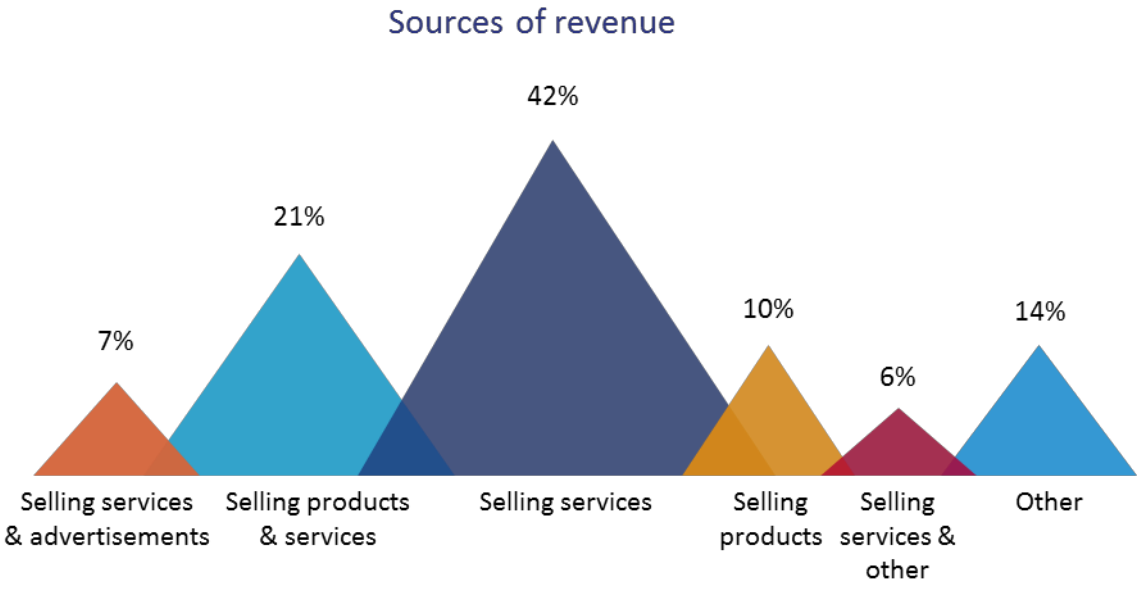


Figure 3: Sources of revenue for the organisations working with Open Data (in percentages)

The services offered by the organisations are not limited to certain domains or groups, as the organisations serve either businesses, governments, consumers or a combination of all client segments. Being digital natives, the organisations distribute their services predominantly using digital channels such as a web shops for applications. Many of the organisations are still in a start-up phase, bootstrapping their organisation to fine-tune their product and conquer markets. Turnover is still only marginally related to Open Data, but this is expected to increase rapidly in the two years to come. However, large discrepancies can be noted across the different companies. Where 23% of the organisations expects their turnover related to Open Data to grow between 11% and 60% in the coming years, 37% of the respondents forecasts their turnover related to Open Data to increase with at least 61% annually.

Although the organisations deliver services in all the steps of the Open Data Value Chain, most of the respondents classified themselves as an aggregator, working on the collection and aggregation of data. This is a counter-intuitive outcome as the study expected to find primarily organisations who use Open Data as a resource for an application, as show-cased most often. This insight shows the increase in the development of the infomediary sector, acting as a data broker. Indeed, Open Data cannot always be directly translated into useful insights, applications or services. It needs to be collected and cleaned first, offering business opportunities for organisations specialised in this particular type of service provision.

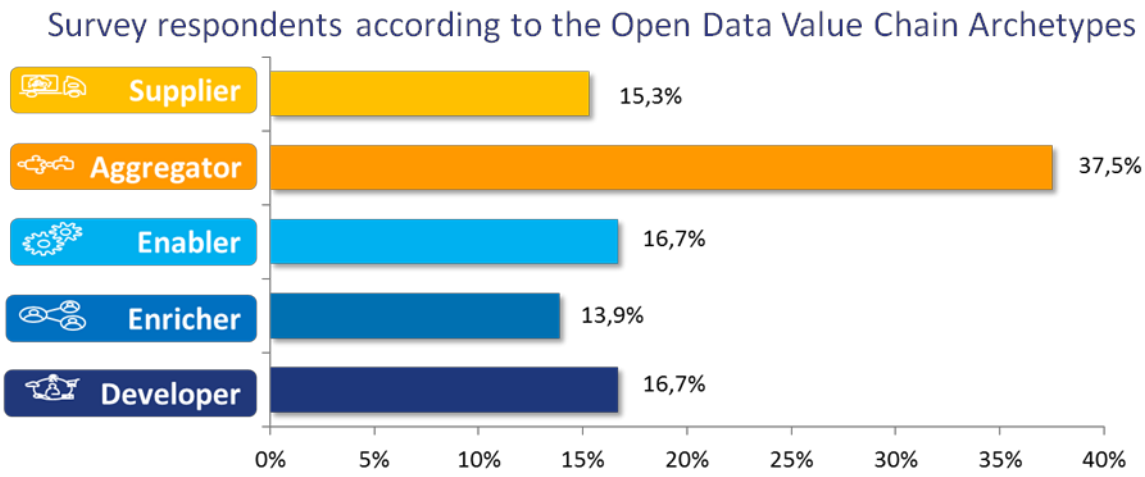


Figure 4: Survey respondents according to the Open Data Value Chain Archetypes (in percentages)

The way in which Open Data is used varies substantially. Examples were found ranging from organisations that use Open Data for the enhancement of their own processes to companies that have built their entire business on Open Data. In particular businesses that have been running for multiple years use Open Data to enhance their existing services or to optimise their processes. Companies also commercialise Open Data within the actual product or service they provide to their client. One way is to make Open Data accessible and workable for others. An example is aggregating all available data belonging to a certain domain into a workable database. The second way to commercialise the re-use of Open Data is to provide insights derived from data analysis in order to help solve the clients' business challenges.

When looking at the people working within these organisations, many of them have a background in ICT and data science (extracting knowledge or insights from data) or strong analytical and statistical skills. This is also the most sought after profile, mentioned by 49% of the organisations (Figure 5). But in order to be successful in transforming Open Data into value, these skills alone are not sufficient. Both hard and soft skills are needed. To create successful and marketable products, it is key to have knowledge and insights from a business perspective as well.

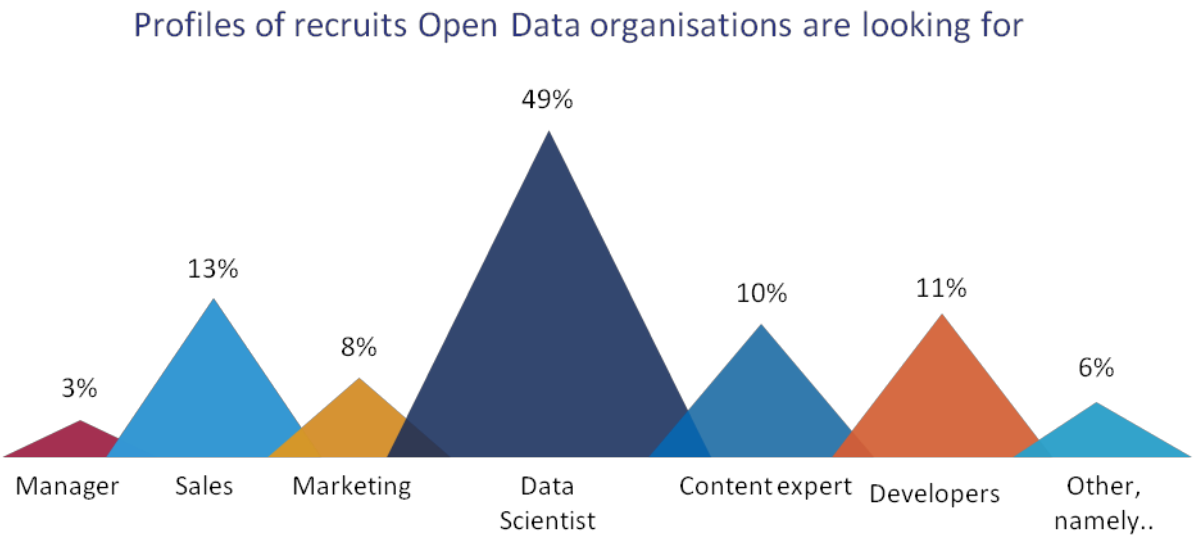


Figure 5: Profiles of recruits Open Data organisations are looking for (in percentages)

Despite the diversity in products and services, the barriers the organisations face when working with Open Data are strikingly similar and can be both internal and external. The external barrier and most frequently mentioned barrier is the quality of the data. At the same time, organisations mention the quality of Open Data as being critical for their operations (Figure 6). Three distinct aspects related to data quality can be identified. The first aspect concerns the content of the data, which might be too old or not granular enough for processing. This also concerns the metadata: it is not always clear whether datasets are updated annually for instance, creating uncertainty for organisations that need to rely on periodically updated data. The second aspect is related to the heterogeneity of the data, such as in format and structure of the data itself. The third aspect concerns the dispersion of the Open Data across numerous portals with numerous platforms, interfaces and languages. These three aspects hinder re-users to standardise or automate the collection and processing of Open Data.

A number of barriers faced when working with Open Data can also be considered as internal barriers. These barriers relate for instance to the skills needed to work with data. Particular skills and tools are needed to fully grasp and exploit Open Data and the low understanding of the benefits it can yield. Following the poor quality of Open Data, many organisations need to spend – disproportionate - time and (financial) effort in the retrieval and refinement of Open Data. Data analysis skills are needed in particular. Finding and allocating people with these skills can be challenging, resulting in many organisations educating their employees 'on the job'.

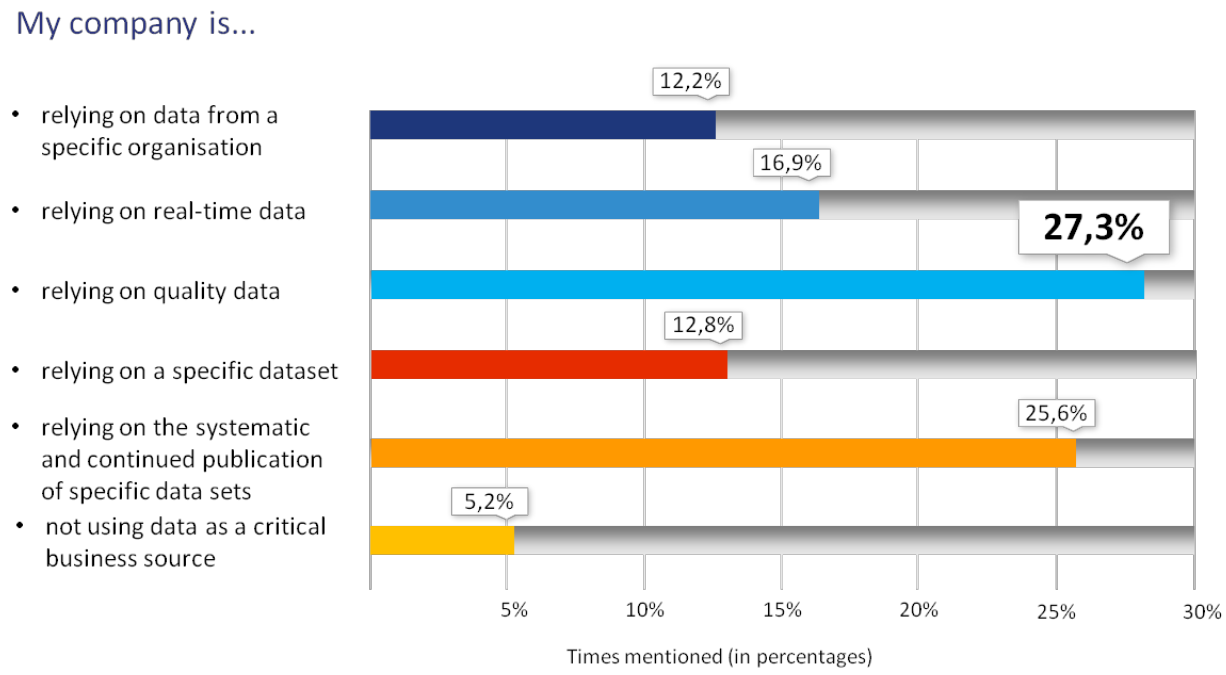


Figure 6: Critical elements for organisations (in percentages).

Following these findings, a number of recommendations can be formulated, addressing both the public and the private sector. For both the public and the private sector, the transcending point of improvement is in raising awareness around Open Data. As indicated by re-users, the current low level of awareness about the potential of Open Data re-use within public administrations results in low quality Open Data. A lack of knowledge on the specific needs of re-users hampers the re-use of Open Data in the private sector. On the other hand, interviewees believe that there is already a goldmine of data available, waiting to be exploited. A better understanding of the potential value in the private

sector could create a virtuous circle in which an increase in the use of Open Data stimulates the provision of more data.

For the private sector to develop further successful business models around Open Data, data needs to be transformed into a product or service that meets the specific business challenges of clients. It is not the development of the product or service which is the most challenging, but creating something that is interesting for the market. When designing a business model relying on Open Data, this should be designed first around addressing a specific challenge faced by potential clients.

Finally, public sector organisations need to better align their data provision strategy with the specific needs of users, resulting in prioritised domains and a better quality of the data and metadata provided. Clear descriptions are needed in order for re-users to be confident that they can discover the data and rely on it over time. Another aspect relates to the heterogeneity of the data, such as in format or (in-) consistency of the metadata, and the dispersion of data among several platforms. These factors hamper automated processes at the side of the re-user. The quickest win in improving the quality of Open Data is by strengthening consistency and standardisation of metadata, preferably using the European standards such the DCAT Application profile and those developed by the ISA Programme of the European Commission. Another quick win can be materialised by centralising the Open Data, either on the national or preferably at a European level.

EUROPEAN DATA PORTAL

TOP 3 SOURCES OF REVENUE OF OPEN DATA COMPANIES

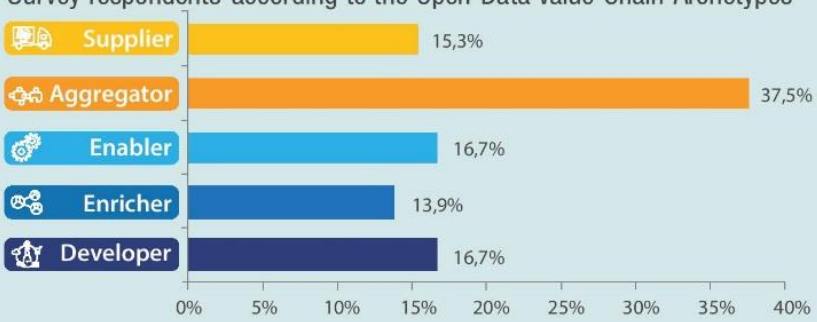
21% Selling products & services

42% Selling services

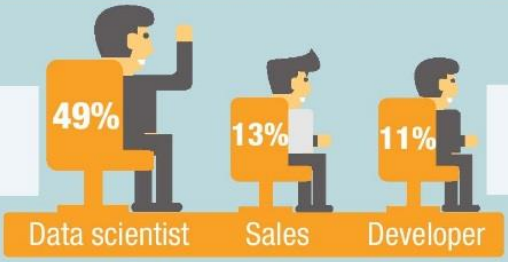
10% Selling products

OPEN DATA ARCHETYPES

Survey respondents according to the Open Data Value Chain Archetypes



TOP 3 PROFILES OPEN DATA COMPANIES ARE LOOKING FOR



ESSENTIAL CONDITIONS FOR RE-USING OPEN DATA

- High quality Open Data
- Systematic and continued publication of Open Data

?

76% OF THE ORGANISATIONS USING OPEN DATA FORESEE TO RECRUIT NEW EMPLOYEES

TAKING OPEN DATA TO THE NEXT LEVEL

- Design your Open Data provision strategy based on user demands
- Standardise and harmonise your Open Data
- Share your story on the use of Open Data

1. Introduction

One of the key requirements to create a healthy Open Data ecosystem in the European Digital Single Market is to ensure smooth access to data across all countries. For this reason, at the end of 2015, the European Commission launched the European Data Portal offering a central access point to Open Data published by national public bodies in countries across Europe. The European Data Portal further monitors the development of national Open Data policies and portals in Europe including the expected political, social and economic impact of Open Data within European countries.

With a 28.6% increase compared to 2015, the EU28+ countries (European Member States plus Iceland, Liechtenstein, Norway and Switzerland, commonly referred to as the EU28+) completed over 55% of their Open Data journey. This shows that, by 2016, a majority of the EU28+ countries have successfully developed a basic approach to address Open Data. The Portal Maturity level increased by 22.6 percentage points from 41.7% to 64.3% thanks to the development of more advanced features on country data portals. Building on these results, to assess the overall Open Data Maturity, the EU28+ are grouped into different maturity levels: Beginners, Followers, Fast Trackers and Trend Setters. Whereas beginners are still struggling with the basics, followers have successfully developed Open Data policies but face challenges in improving data accessibility. Fast trackers have significantly accelerated their Open Data journey over a short time span, having either a policy or a portal that is substantially developed, however, they still face a small number of shortcomings in reaping the full benefits of either their policy or portal. Trend Setters on the other hand, have implemented an advanced Open Data policy with extensive portal features and national coordination mechanisms across domains. The groupings of European countries are illustrated by Figure 7.[4]

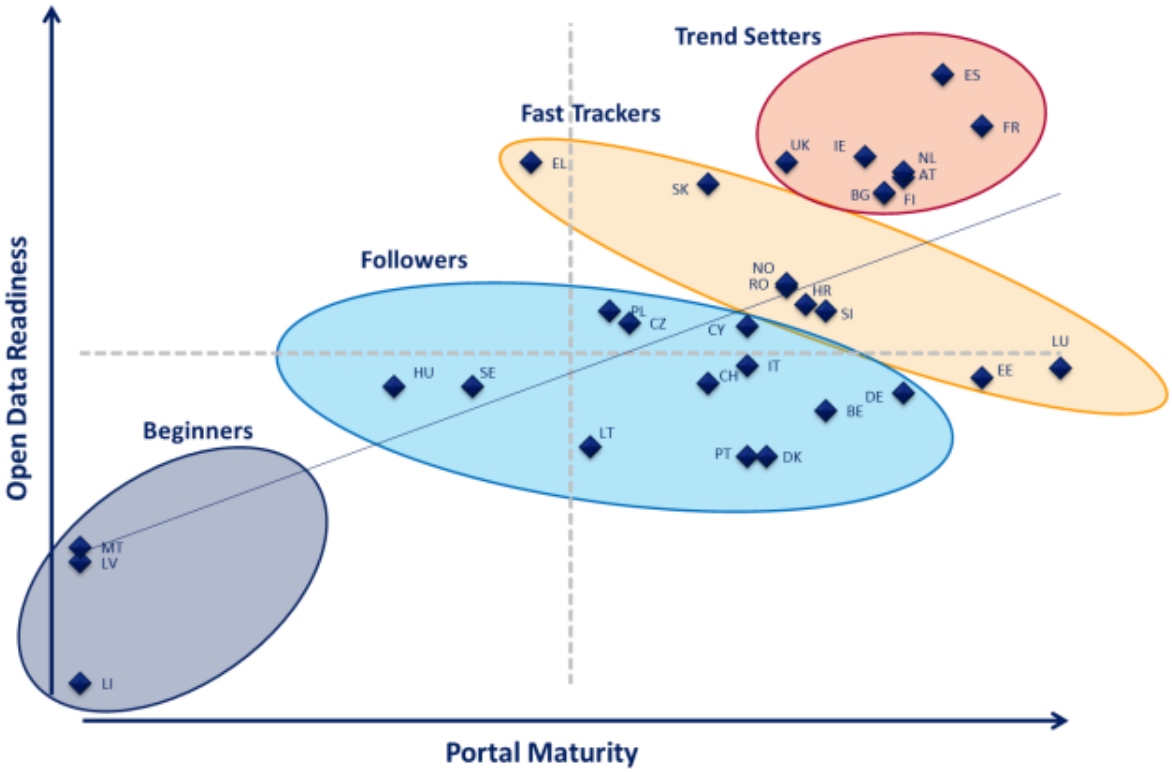


Figure 7: Open Data Maturity Clusters

[4] European Data Portal, 2016, Open Data Maturity in Europe

Barriers do remain to move Open Data forward. Among others, countries need to move forward with the effective implementation of their data policies, emphasising the importance of a legal structure addressing licensing, privacy aspects and standards. Further efforts should be invested in improving data quality, including the development of automated processes which will increase the usability of data, and the organisation of more events and trainings to support both local and national initiatives.

Studies measuring the impact of Open Data initiatives underline the importance of Open Data for economic growth, to drive monetary benefits and foster transparency. Especially since Open Data generally has a marginal cost of zero, further economic benefits are likely to result from more organisations re-using Open Data. In this context, Open Data – as defined by the Open Definition - refers to the information collected, produced or paid for by public bodies and can be freely used, modified, and shared by anyone for any purpose. Furthermore, data must be available under an open licence and preferably in machine readable format. This definition can be narrowed down to Public Sector Information (PSI), or Open (Government) Data, that is to say data collected and published by the public sector.

Whereas the macroeconomic impact of Open Data is undeniable, further understanding of how it is used at company level has been on the wish list of many data publishing organisations. The aim of this study is therefore to present examples of the economic impact of Open Data at the microeconomic level by focusing on individual organisations and primarily companies that use Open Data to generate further (economic) value.

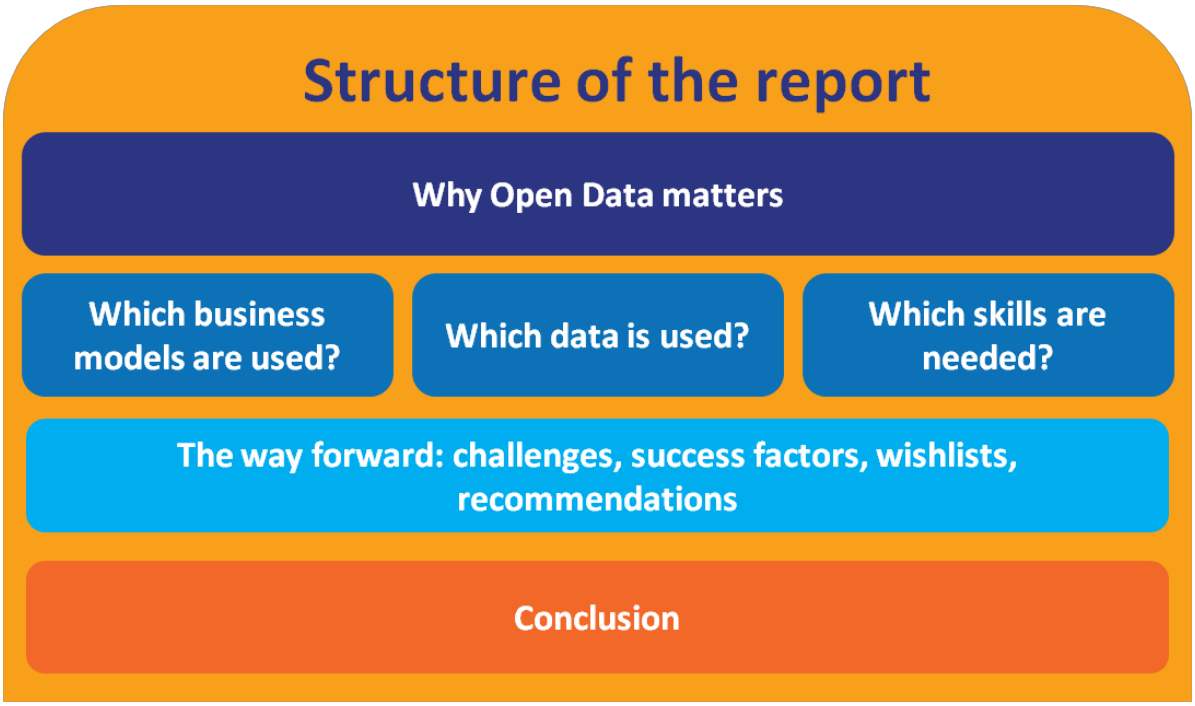


Figure 8: Structure of the report

Following the present introduction, Chapter 2 summarises the economic value of Open Data at the macroeconomic level to underline the overall economic and societal benefits of Open Data. It concludes by presenting the method used to understand what businesses are doing with Open Data.

The following chapters of the report provide further in-depth insights into what the individual organisations do with Open Data. Chapter 3 explores what types of data are being reused the most, where and how they are accessed and what benefits different organisations witness in working with

Open Data. Understanding how organisations generate revenue from the re-use of Open Data is also key in understanding these benefits (Chapter 4) and which skills are needed to work with Open Data (Chapter 5). Chapter 6 assesses the key challenges and success factors in working with Open Data and provides a series of recommendations for both the public and private sector on how to increase the economic value generated by the use of Open Data. Finally, a conclusion is provided in Chapter 7.

- Annex I offers a comprehensive overview of the methodology used to outline the findings presented in the report.
- Annex II contains the list of organisations interviewed
- Annex III presents a correlation matrix showing data re-use across categories
- Annex IV is the interview questionnaire
- Annex V is the survey questionnaire
- Annex VI is the survey used for the temperature check conducted during IODC 2016.

2. Why Open Data Matters

Increasing amounts of Open Data are made available on the Internet, by a number of organisations. By harvesting close to the exact same sources in 2016, the European Data Portal has more than doubled the amount of data it references. Indeed, over 600,000 datasets are now referenced compared to 240,000 when the portal was launched over a year ago in November 2015. However it is not the sheer amount of data that drives the impact. It is the potential encapsulated by the data itself. One way of assessing the impact of Open Data is to monetise its value.

To assess the economic value of Open Data at a microeconomic level, it is important to first understand what the economic value of Open Data is at the macroeconomic level. Therefore, this chapter provides an overview of the impact of re-use of Public Data Resources for the EU28+, after first having explored the Open Data Value Chain in order to understand the background of this expected economic value. To measure the impact, four key indicators were measured: direct market size in euro and percentage of Gross Domestic Product (GDP), number of jobs created through the availability of Open Data, cost savings that can be realised in the public sector, and efficiency gains for everyone.

2.1 The Open Data Value Chain

The first step in the value chain is to create data. The data is then validated and released, for example through a portal or bought by a private company, after which it can be analysed. By aggregating different data sets, new data is created which can lead to new data services or products. Finally, these data services and products can be further aggregated. This process is visualised below in Figure 9.

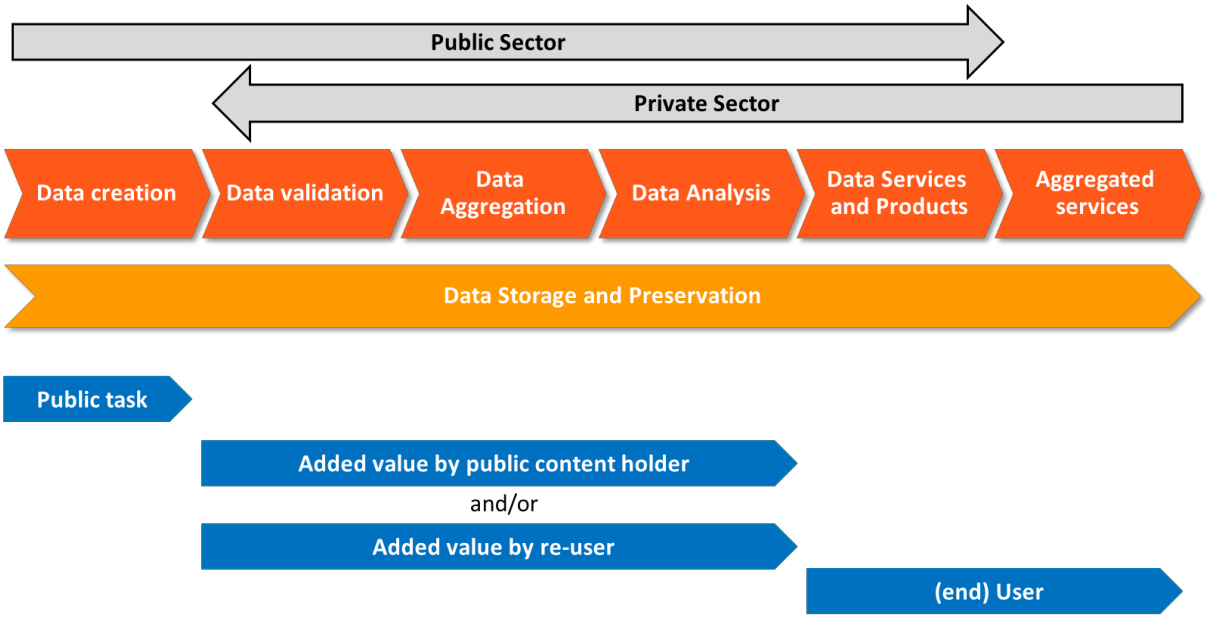


Figure 9: Open Data Value Chain

Various actors are involved in this process. Following the steps above, data creation is done by the suppliers. The data is subsequently collected and aggregated by the so-called aggregators. Two types of archetypes are then identified who use the data in their work, developers and enrichers. These archetypes either use the data for the development of new applications (developers) or to gain new and better insights from analysis of the data (enricher). These new insights could often only be achieved through the use of Open Data and are therefore essential in the Open Data value chain. Enablers facilitate the supply or use of Open Data for the other archetypes. By providing platforms, on

the one hand they generate revenue, while on the other hand they provide cost-effective and easy-to-access services for both data suppliers and data consumers. The different archetypes therefore all add value to the Open Data at different steps along the value chain. This is further illustrated in Figure 10.

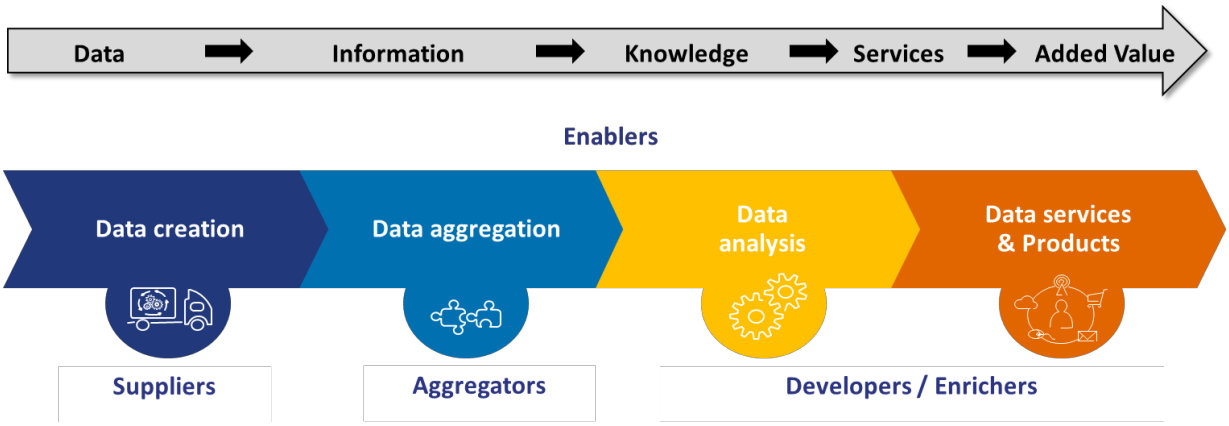


Figure 10: Data Value Chain Archetypes

By making more data freely available for re-use, both public bodies and the private sector can create data services and goods, as well as aggregated services further downstream. Different services are then purchased across the data value chain, leading to enhanced end products and services that can either be purchased or delivered free of charge to an end client.

2.2 The Economic Impact of Open Data

Government data made available as Open Data can either have direct or indirect benefits. By identifying these benefits, governments are better equipped to measure the success of their Open Data policies. In this respect, direct benefits are considered to be monetised benefits that are realised in market transactions in the form of revenues and Gross Value Added (GVA), the number of jobs involved in producing a service or product and cost savings. Indirect economic benefits are for example new goods and services, time savings for users of applications using Open Data, increased efficiency in public services and growth of related markets.

To create innovative products, it is essential for data to be freely accessible and to be in the appropriate format for re-use. The value of each data source can be further increased by combining multiple data sets to create a bigger playground. This can lead to more competition which in turn leads to further innovation and use of Open Data. In addition to publishing more data, it is also important to incentivise people to make use of Open Data. For this to happen, both governance mechanisms and a clear understanding of the user's perspective are needed. Feedback mechanisms provide an excellent opportunity to create this interaction between governments and individuals. Also companies have increasingly started to use Open Data by entering the Open Data Value Chain. This in turn leads to more possibilities of re-using Open Data and more data sets having an economic impact. The following sections will estimate the value created through Open Data in the EU28+ looking at both financial and societal benefits.

2.2.1 Value created through Open Data in the EU28+

The study 'Creating value through Open Data'[5] uses four key indicators to quantify the potential size of the Open Data market in the EU28+, being direct market size, number of jobs created, cost savings and efficiency gains. For 2017, the direct market size of Open Data is expected to be 59.7 bn EUR for

[5] European Data Portal, 2015, Creating Value through Open Data

the EU28+. The market size is expected to increase by 36.9% between 2016 and 2020, up to a value of 75.7 bn EUR in 2020. The cumulative direct market size between 2016 and 2020 is estimated at 325 bn EUR. Considering both the direct and indirect market size of Open Data, the cumulative total market size for Open Data is estimated between 1,138 and 1,229 bn EUR. The direct market size differs per country in the EU28+. Generally speaking, the market for Open Data still needs to be developed in Eastern European countries, the Baltics and Iceland. On the contrary, in countries such as France, Germany, Spain and the United Kingdom the market for Open Data is already significant (Figure 11).

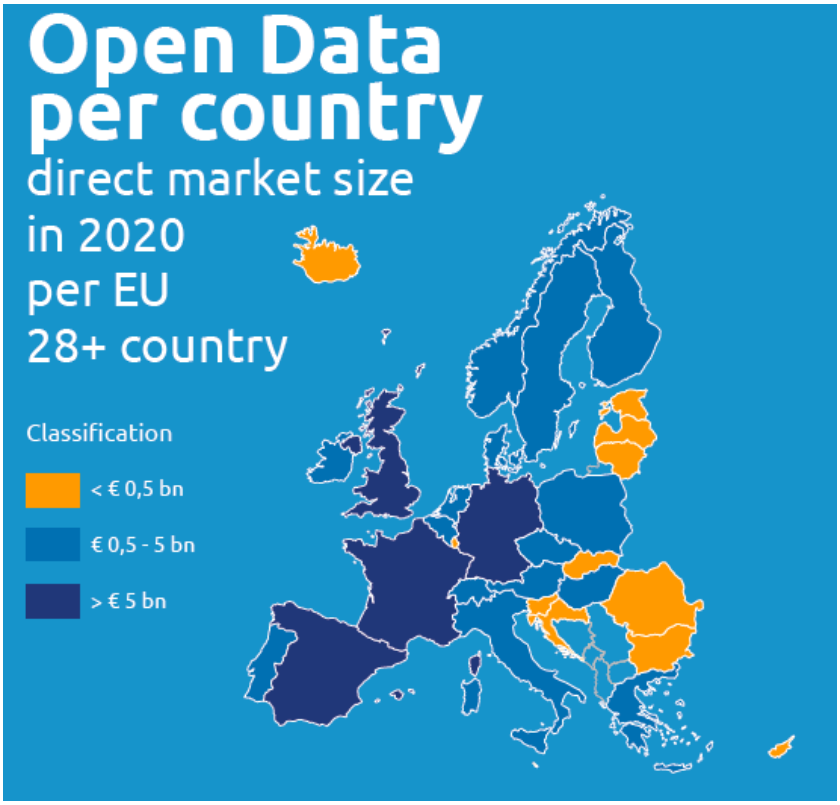


Figure 11: Direct market size of Open Data per country[6]

In terms of market share as percentage of GDP, the GDP is estimated to grow from 2016 to 2020 with the forecasted inflation figures only. This means an increase of 7.7% in 2020 compared to 2015 estimating GDP in 2020 to reach 15,998 bn EUR. The market share of Open Data as a percentage of GDP is therefore expected to have increased to 0.47% by 2020. The expected increase of the Open Data market share as a percentage of GDP from 2016-2020 is shown in Figure 12.

	2017	2018	2019	2020
Share of GDP	0.39%	0.42%	0.44%	0.47%

Figure 12: Forecasted share of direct Open Data in EU 28+ GDP, 2016-2020

In terms of market size per sector, public administration is the sector to gain the most from opening up data and thereby being the first and most important re-user of (its own) Open Data (Figure 13). At the same time, agriculture, arts and the entertainment sectors are expected to gain least from Open Data. This does not mean that Open Data has little impact in these sectors, it merely shows that the estimated impact in this sector will take more time to reach its full potential.

[6] Ibid

Figure 13: Total market size of Open Data per market sector for EU28+ in millions, 2020

The second indicator aimed at finding out how many jobs would be created in the EU28+ private sector that could be directly linked to the re-use of Open Data. It was estimated that in 2017, 80.500 jobs would be directly linked to the re-use of Open Data. Since the market size of Open Data is expected to increase quite substantially by 36.9%, an extra 25,000 jobs directly related to Open Data are expected to be created by 2020, showing an average growth rate of 7.3%. This is shown in Figure 14.

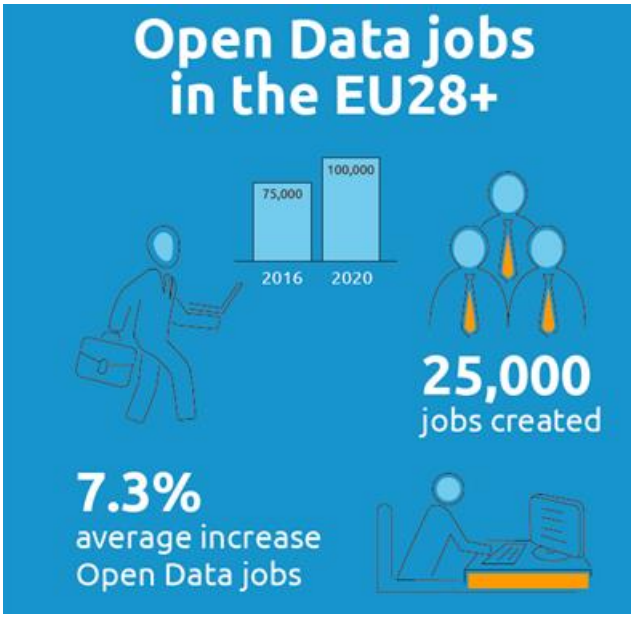


Figure 14: Forecasted Total number of direct Open Data jobs (in person), 2016- 2020

The third indicator focused on obtaining a better understanding of the impact of Open Data on cost savings within public administrations. The study used the Danish government calculation to estimate the cost savings percentage for all countries. By taking into account the forecasted 2020 GDP as well as the respective government expenditure averages, the cost savings percentage per country could be calculated. In total, the cost savings for the EU28+ in 2020 were forecasted to be 1.7 billion EUR.

The fourth and final indicator looked at the non-economic impact of Open Data, by focusing on minimising waste and maximising the outcome value by improving resource allocation. This non-economic indicator is further explored in the next paragraph.

2.2.2 Societal benefits of Open Data

The economic benefits resulting from the use of Open Data are substantial, but there are numerous positive effects which cannot be financially quantified. Openly sharing data can reduce search costs, makes monitoring easier, enables faster and easier access to information, better resource allocation, and increases automation, standardisation and interoperability. In the 2015 report 'Creating value through Open Data'[7], the metric 'efficiency' was used in order to assess the benefits realised through efficiency gains. The aim of efficiency is to improve resource allocation so that waste is minimised and the outcome value is maximised, given the same amount of resources. Three exemplar sub-indicators are assessed in more detail: how Open Data can save lives, how it can be used to save time in the transportation sector and how Open Data helps achieve environmental benefits.

How can Open Data save lives? Open Data not only helps patients choosing a healthcare provider based on treatment outcomes, but also stimulates healthcare providers to share their best practices. In emergency situations where every minute counts, Open Data can improve an effective response by analysing where to place equipment and station personnel. It was estimated that Open Data has the potential of saving 1,425 lives a year equalling 5,5% of the European road fatalities.

Secondly, Open Data saves time in the transportation sector. It is important to focus on reducing travel time, as congestion costs in Europe represent 1% of GDP every year. Providing people with an alternative route based on real-time data increases the distribution on different roads and reduces the length of traffic jams. In terms of time saving, it was estimated that when Open Data is applied to the transportation sector, a reduction of 10% waiting time would save a total of 629 million hours of unnecessary waiting time on the road in the EU between 2016 and 2020.

Figure 15: Examples of non-economic benefits of Open Data

Thirdly, Open Data helps to achieve environmental benefits. More and more, governments are searching for new ways of reducing the toxic effects of, for example, CO_2 emission and improving waste management. Open Data can be useful to reduce those adverse effects as well, by providing more insight into the specific areas where those problems cause the highest health risks and act on them.

These are just three examples, containing only a small selection of the enormous amount of useful applications that exist. The societal benefits of Open Data are numerous. Social welfare can be improved as society benefits from the transparency accessible information creates. Open Data enhances collaboration, participation and social innovation. The availability of public Open Data not only brings information to citizens, but also enables them to engage in decision-making processes. It empowers citizens to become agents of social transformation by monitoring and overseeing government actions and public policies.

[7] European Data Portal, 2015, Creating Value through Open Data

2.3 Understanding the benefits of Open Data at company level

Open Data offers a substantial economic and societal potential. However, this can only be realised when the data is accessible, re-usable and actually re-used by organisations transforming it into value. In order to exploit this potential, it is therefore important to gain insight into the economic and societal value and experiences at company level. In other words, delving from the macroeconomic benefits to the microeconomic level. What are the experiences of companies when accessing or re-using data? What barriers are they facing? How much money do they generate from the re-use of Open Data? To better understand these companies, a three-pillar approach has been have developed which is displayed in Figure 16. This section highlights the core components of each of the three pillars, full descriptions are available in Annex I.

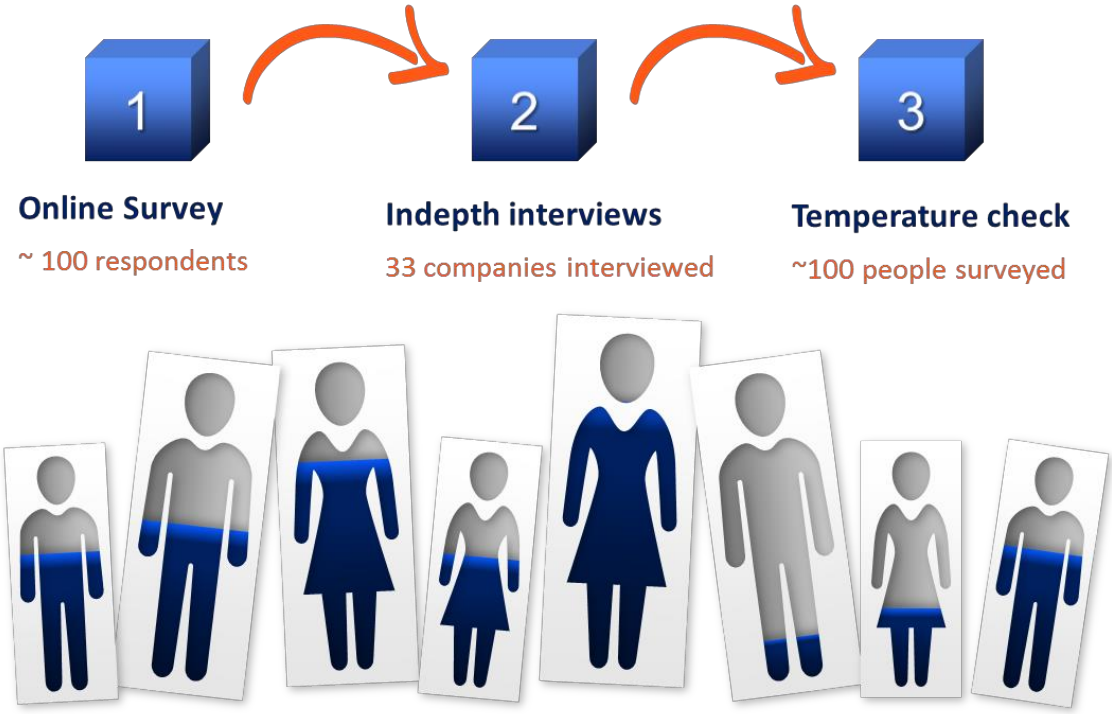


Figure 16: Three Pillar Approach

First, a survey was conducted targeting businesses. The survey consisted of 57 questions, most of which were multiple choice. The survey was open to the public from 18/02/2016 and ran until 05/06/2018 and was made accessible on the European Data Portal. The full questionnaire can be found in Annex VII. Close to a 100 answers were collected, of which 76 complete questionnaires could be used within the assessment.

After having excluded the incomplete entries, the database consisted of 76 entries from 21 countries, six of which are non-European.[8] In the analysis of the survey results, the non-European countries are included in the calculations. Unfortunately, not all EU Member States are represented by the respondents and within the sample a limited number of Member States are over-represented. This has led to a disproportionate representation of certain segments and has to be taken into account when trends are identified in the results. The country distribution of the valid respondents can be seen in Figure 17.

[8] Two entries from Canada and one from respectively Chile, Australia, Nigeria and Turkey

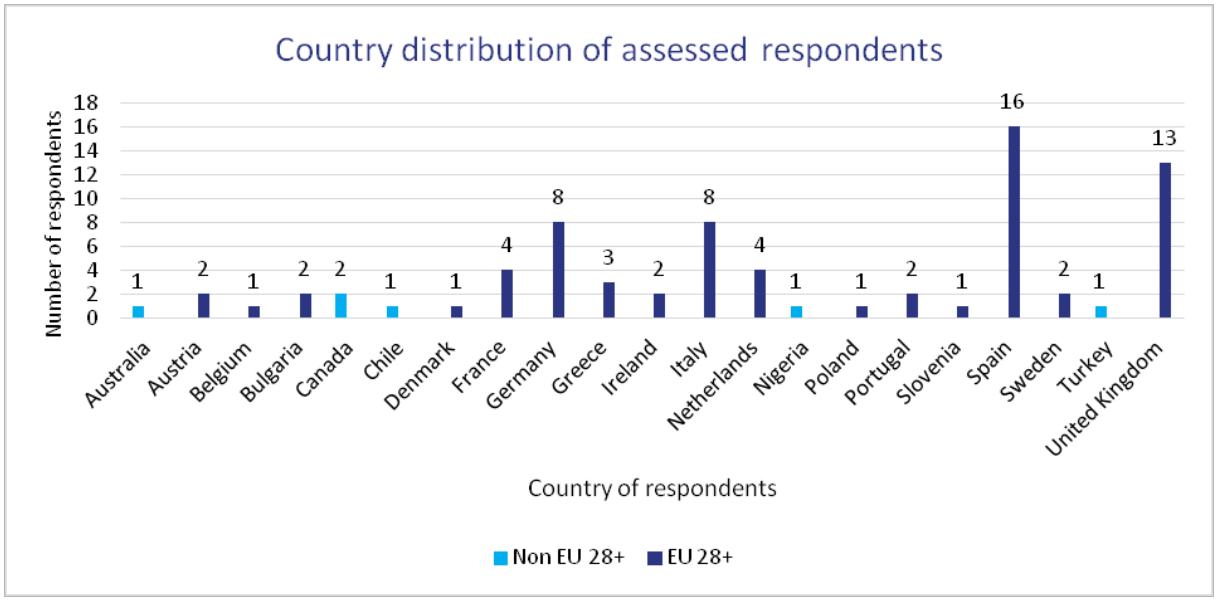


Figure 17: Country distribution of assessed respondents (in absolute numbers)

Of the valid entries, fifteen organisations indicated to be active in more than one location. While the majority of these are active in one other location, the others are active in up to six locations. Interesting to see was that four respondents stated to be active not in particular countries but in (geographical) regions or groupings, for example the European Union or in Extractive Industries Transparency Initiative (EITI) countries.

Secondly, in-depth interviews were conducted with 33 organisations in order to gain further insight into their experiences in working with Open Data. A solid geographical spread across the European Union and a spread across different NACE codes were important criteria when selecting the organisations to be interviewed. Via the interviews, common challenges were identified as well as the diversity in possible business models with Open Data became apparent. Together with the main results from the survey, the 33 organisations interviewed are presented in greater detail throughout the next chapters. A table listing the companies interviewed is available at the end of this chapter.

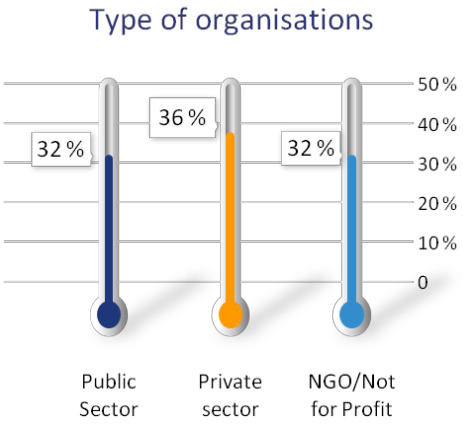


Figure 18: Type of organisation – IODC Temperature Check (in percentages)

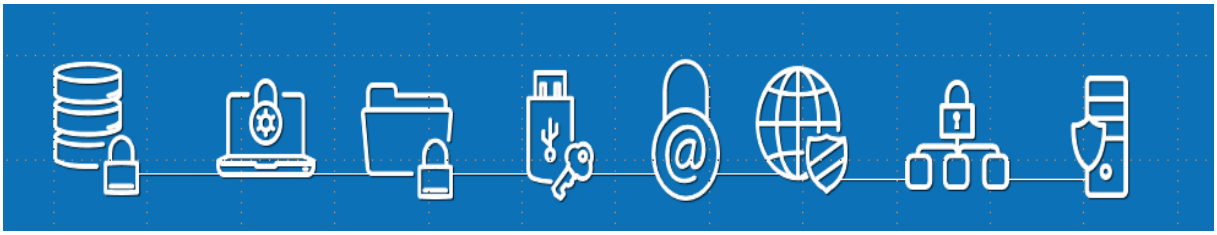
Thirdly, a temperature check was conducted during IODC 2016. The results of the survey and the interviews are complemented by the results of the 'IODC Temperature Check': a short survey filled out by a sample of participants of the International Open Data Conference, on 6-7 October 2016 in Madrid. A total of 92 organisations filled in a set of 5 questions relating to their experiences with Open Data, with a specific focus on the main benefits, challenges and type of data used. Since most organisations assessed in this report are from the private sector, grouping the answers of the temperature check by type of organisation provides the best opportunity to cross-check the findings of the survey and understand the different and similarities in perceiving the re-use of Open Data. 32% of the respondents are from the public sector, 36% from the private sector and 32% from the non-government/ non for profit sector. The spread of the respondents is shown in Figure 18.

The next chapters will dive deeper into the economic impact of Open Data at the microeconomic level by presenting and analysing the results of this study. The table below lists the companies interviewed during the second stage of this study. The companies are presented more in-depth in Annex II: List of organisations showcased.

Organisation name	Website	Country
Appchallenge	www.appchallenge.net/	United Kingdom
BBVA	www.bbvadata.com/	Spain
Canard Drones	www.canarddrones.com/	Spain
Connemara Programme	www.myconnemara.com	Ireland
Crige PACA	www.crige-paca.org/	France
CropDiagnosis	www.cropdiagnosis.com	Greece
Doctrine.fr	www.doctrine.fr	France
Egovlab	www.egovlab.eu	Sweden
eVineyard	www.evineyardapp.com	Slovenia
Fluicity	www.flui.city	France
Geomanagement	www.levadia.gr	Greece
Glimworm Information Technologies	www.glimworm.com	The Netherlands
GraphDefined GmbH	www.graphdefined.de	Germany
Hammer Project	www.hammer-project.com	Italy
Infoempresa.com	www.infoempresa.com	Spain
ISEA	www.iseamcc.net	Spain
Kenedict Innovation Analysis	www.kenedict.com	The Netherlands
Mozaika	www.mozajka.co	Bulgaria
Netage	netage.nl/en	The Netherlands
Normative	www.normative.io	Sweden
OECON GROUP	www.oecon.gr/en	Greece
Openlaws	openlaws.com/	Austria
OpenMove	openmove.com	Italy
Open Opps	www.openopps.com	United Kingdom
Peacemeetings	menplanner.org	Spain
SmartAppCity	www.smartappcity.com	Spain
Smartbow	www.smartbow.at	Austria
Smartup Cities	www.smartupcities.com	Spain
UniGraph	unigraph.rocks	United Kingdom
Wikimedia	www.wikimedia.org	Belgium
Wise Town	wise.town	Italy
Wripl technologies	wripl.com	Ireland
Zimmerman & Zimmerman	www.zimmermanzimmerman.nl	The Netherlands

3. What data do businesses re-use?

Before delving into the business models developed to generate revenue from Open Data re-use, it is important to first understand what it is that organisations are doing with data. This chapter will assess the types of data that are the most re-used, where the data is found, how it is accessed, data formats and finally the benefits seen by data re-users. To back the findings of this chapter, the answers provided by 76 organisations were assessed.



3.1 Which data categories are the most re-used?

The objective of this study is to understand how organisations, primarily within the private sector, use Open Data and how they have sucessfully developed business models around its re-use. When looking at the spread of data categories, comparisons between companies can be made. Data categories are categorisations of data sets linked to a common theme. An example is wind turbines in coastal and maritime areas in Sweden. This data set can be categorised as part of the environment data category.

To understand which data categories are being re-used the most, respondents could select a total of 13 categories as identified by the European Commission in the context of the revision of the European DCAT Application Profile[9]. These 13 categories were identified on the basis of Eurovoc domains[10], categorisation of data done by Eurostat and comparisons across a number of European countries and other international organisations. As Eurovoc is a multilingual thesaurus, it offers the benefit of mapping terms from all 24 official EU languages.

All 76 respondents indicated to re-use at least one or more data categories, spreading across all 13 data categories identified. The top-five most used data categories are government & public sector (11.9%), economy & finance (11.6%), regions & cities (10.1%), population & society (9.5%), and environment (8.9%). These five data categories together represent 52% of the total re-use of Open Data by the respondents. The least used data categories are international issues (3.6%), health (4.2%) and justice, legal system & public safety (4.2%). Figure 19 below illustrates the spread of re-use across data categories by Open Data organisations.

[9] Joinup, 2016, DCAT-AP

[10] http://eurovoc.europa.eu/

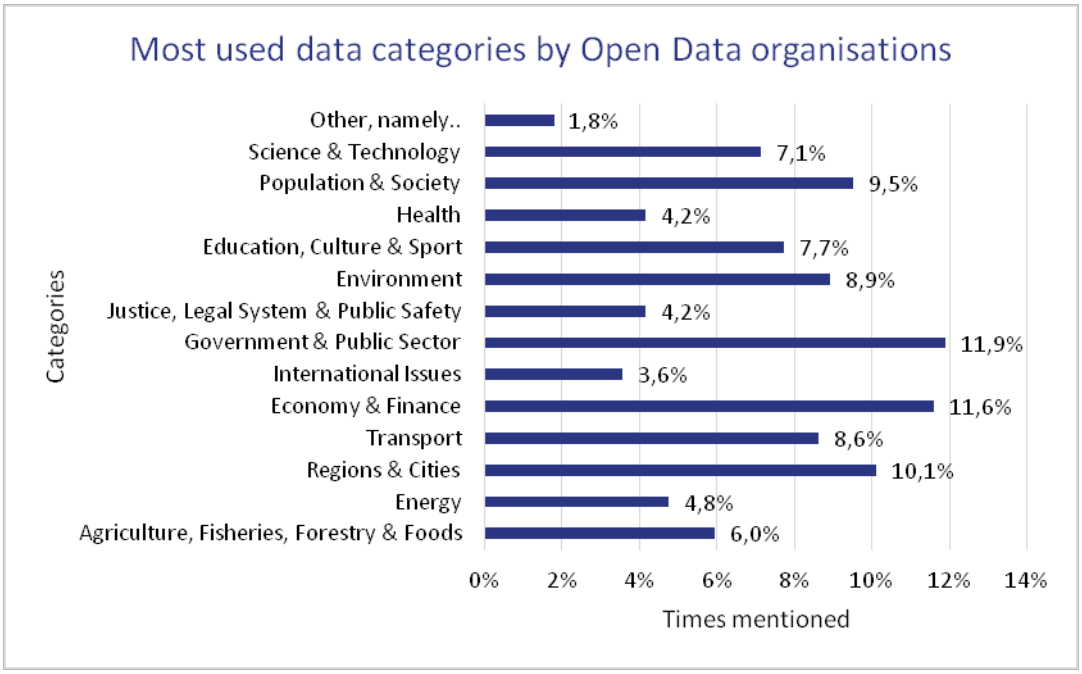


Figure 19: Most used data categories by Open Data organisations (in percentages)

The wide usage of different types of data is also reflected by the number of organisations that indicated to use data from more than one category. Although fifteen organisations only make use of one type of data,[11] on average the organisations make use of data from nearly five different data categories.

To examine the correlation between the use of different data from different categories, the categories were plotted in a correlation matrix which can be found in Annex III. This matrix displays the likelihood of each category to be selected together with a different category. The most noticeable correlations were found between transport and population & society (68% chance), transport and regions & cities (60% chance), transport and environment (52% chance) and population & society and environment (51% chance). Categories with the lowest chance to be selected with another category are government & public sector and education, sport & culture (2% chance), or government & public sector combined with science & technology (2% chance).

Comparing the most popular data categories to the datasets available on the European Data Portal (EDP) results in interesting findings. The EDP data category offering most mapped data sets is the justice, legal system & public safety category (27.8%), followed by environment (23.6%), regions & cities (12.0%), science & technology (11.9%) and population & society (5.5%). The category government & public sector provides only 3.6% of the total mapped data sets while the economy & finance category only provides 4.4%. A solution for this mismatch between available data sets and re-used data could be for public administrations to develop a publication strategy which is more aligned to the needs of data users. An overview of this comparison is provided in Figure 20 below.

[11] Of which the data category "Education, Culture & Sport" is the most popular one with three respondents

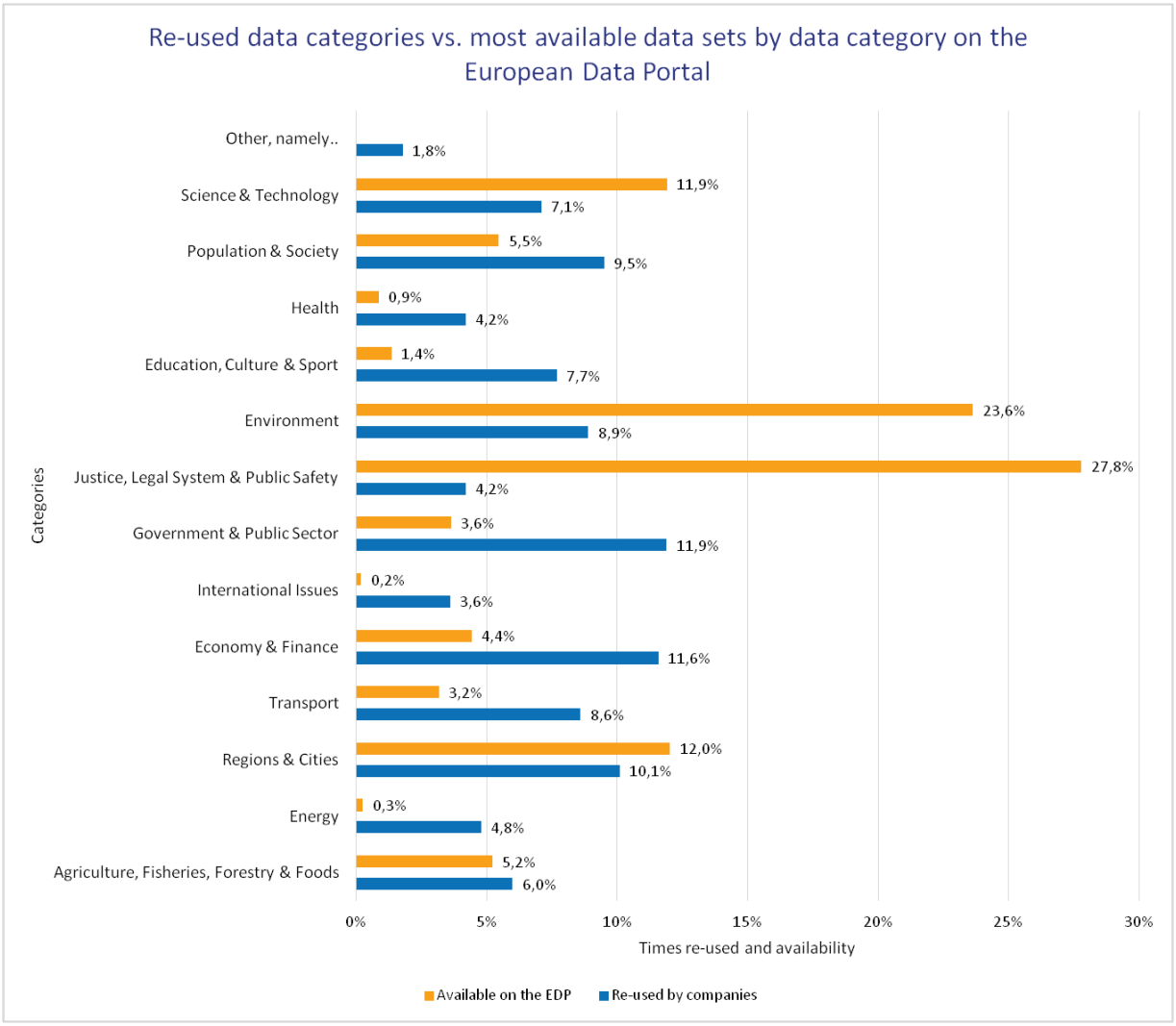


Figure 20: Re-used data categories vs. most available data sets by data category on the European Data Portal (in percentages)

The data re-used by the respondents to the first survey can also be compared to a series of high priority data domains. These domains are the five thematic data domains identified by the European Commission that are expected to represent those with the highest demand from re-users across the EU[12] (Figure 21). These datasets should therefore be prioritised and be made available by the various (public) data providers.

[12] European Commission, 2014, Guidelines on recommended standard licences, datasets and charging for the reuse of documents

	Domain	Examples of datasets
1.	*Geospatial data*	Postcodes, national and local maps (cadastral, topographic, marine, administrative boundaries, etc.)
2.	*Earth observation and Environment*	Space and in situ data (monitoring of weather, land and water quality, energy consumption, emission levels, etc.)
3.	*Transport data*	Public transport timetables (all modes of transport) at national, regional and local levels, road works, traffic information, etc. (*).
4.	*Statistics*	National, regional and local statistical data with main demographic and economic indicators (GDP, age, health, unemployment, income, education, etc.)
5.	*Companies*	Company and business registers (lists of registered companies, ownership and management data, registration identifiers, balance sheets, etc.)

Figure 21: High priority data domains

Respondents were asked from which of these five priority domains they mainly re-used data. Thirteen out of 76 respondents (17%) indicated not to use any of the above. The remaining 63 respondents (83%) indicated to re-use data from approximately two priority domains, resulting in a cumulative total of 128 mentioned domains. As can be seen in Figure 22 most re-used data appears to be statistical data (27.3%) followed by geospatial data (25.8%), companies (19.5%) and transport and infrastructure (18.8%). Only 11 respondents (8.6%) indicated to use data from the domain "earth observation".

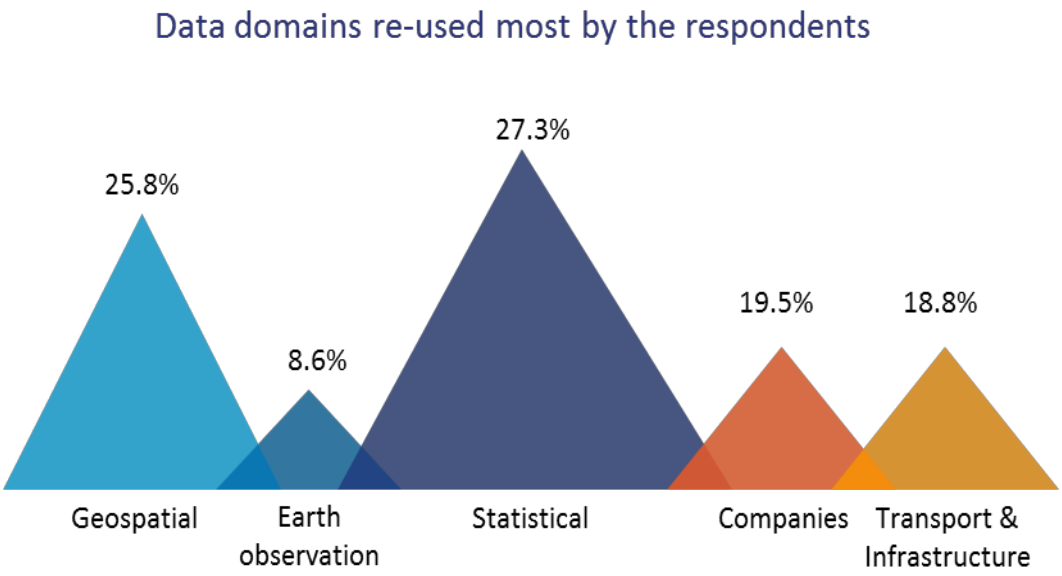


Figure 22: Data domains re-used most by respondents (in percentages)

The five most re-used data domains as indicated by the survey respondents overlap to a large extent with the five data domains consulted most on national Open Data portals generally, as shown in Figure 23. In both surveys, statistical data and geospatial data are most consulted and re-used while earth observation seems to be the priority domain that is re-used the least. Where in this survey company data was re-used more than transport data, the survey conducted in the EU Member States showed that transport data was consulted more than company data. However, the difference is marginal. This shows a clear correlation exists between most consulted data domains available on national portals and most re-used data by organisations. At the same time, providing less data sets on national portals on i.e. earth observation and environment than on statistical data could be a reason why less companies re-use data from this data domain.

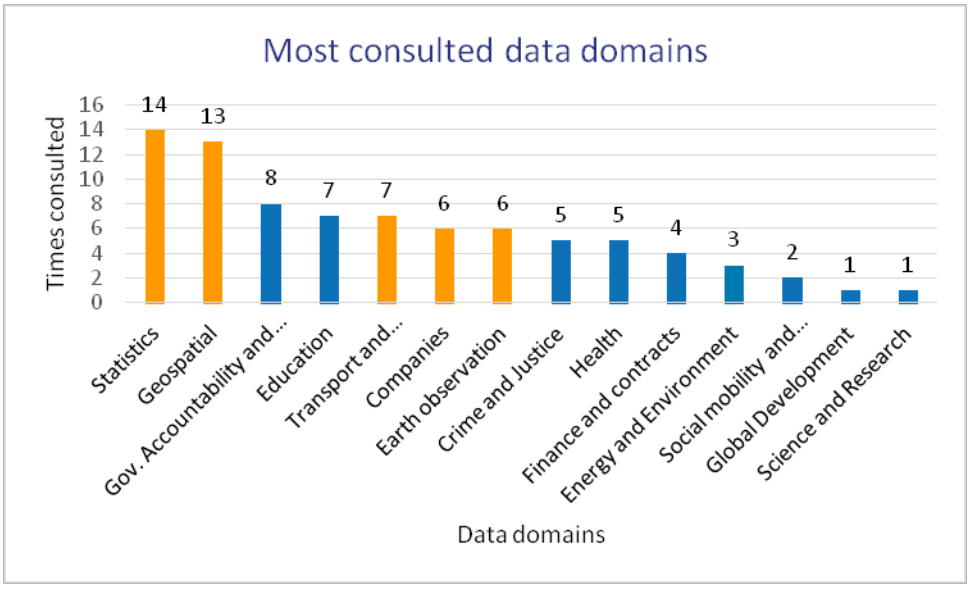


Figure 23: Most consulted data domains by the EU28+ countries (in absolute numbers). High priority data domains appear in orange[13]

To analyse the relations between the most re-used data categories, the preferences of the respondents were plotted in a correlation matrix which is shown in Figure 24 and in Annex III. This matrix shows that most categories have a relatively strong correlation with data from a particular other category. The most noteworthy correlations are found between the category "Regions & Cities" with either "Transport", "Environment" or "Population & Society" which all have a value of over 0.5 thus suggesting a strong trend of organisations using these categories together. This finding correlates with the steady development of smart cities where Open Data and transport and housing are increasingly regarded as one of the most important subjects in the continuous growth of smart cities[14].

From the 76 respondents, 36% of the respondents positioned themselves as data aggregators meaning they do not use the data directly but enable others to publish their data instead. This confirms the fact that a growing number of companies play a role in the infomediary sector, acting as data brokers for their clients.

[13] European Data Portal, 2016, Open Data Maturity in Europe

[14] European Data Portal, 2016, Open Data in Cities

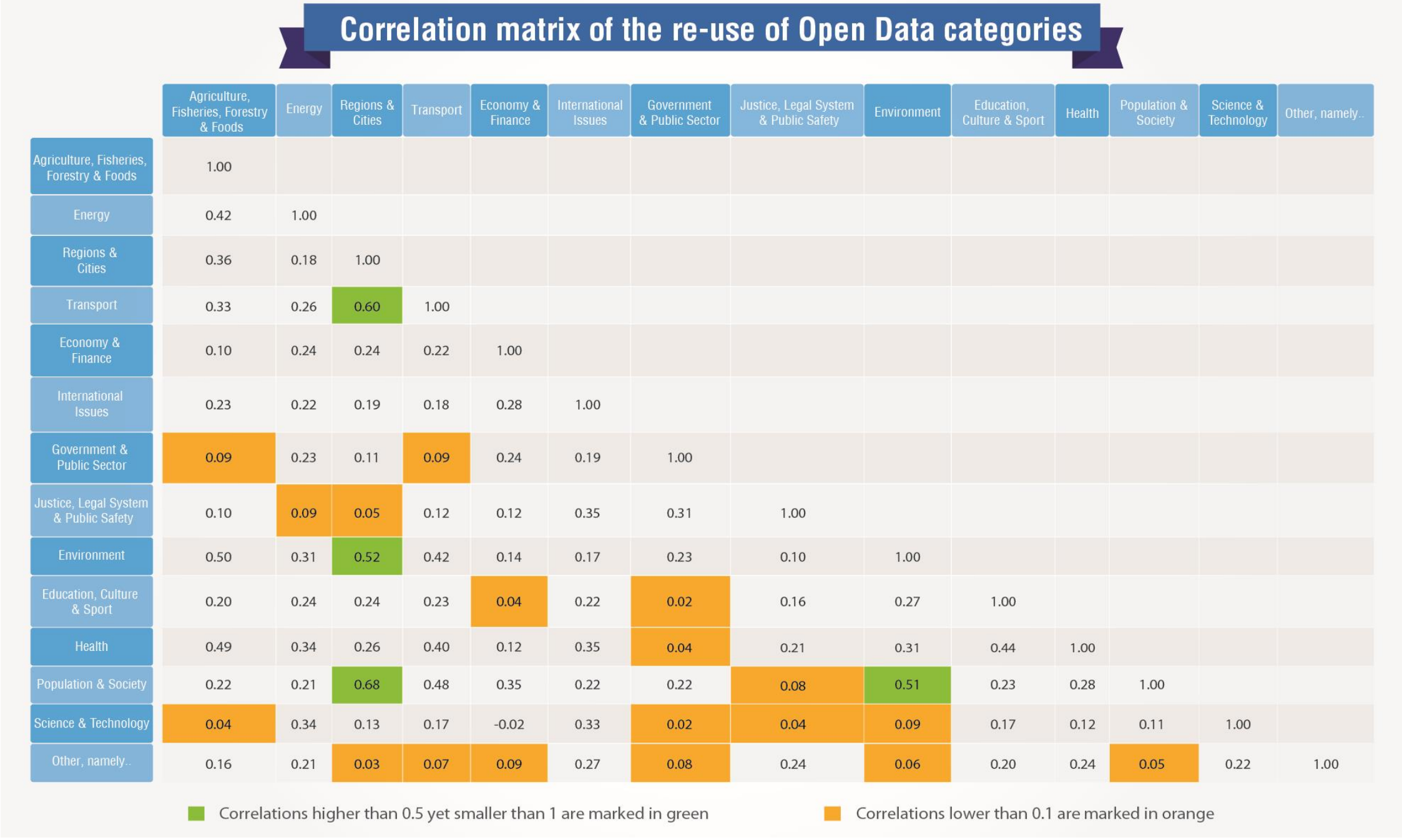


Figure 24: Correlation matrix of the re-use of Open Data categories

Considering the size of the sample assessed it is noteworthy to reference niche players as well. Four organisations use a particular niche of data which is not reflected in the mentioned categories, such as data on development aid. This latter company, Zimmerman & Zimmerman is presented more in depth in Chapter 5.

OpenMove is a good example of a company working in and using data from one particular domain. This Italian company is providing a ticketing solution in the form of a mobile application suite for both commuters and transportation or parking agencies. Via this application, citizens are provided a convenient tool to buy tickets and the agencies are provided means to control their customers using digital tools. For this company, the product was not only the result of their expertise but also related to the availability and quality of Open Data in this particular field. Thanks to the support of Google, many providers have adopted a single standard which improved the quality and consistency of Open Data. This shows the potentially virtuous circle between (quality) Open Data and economic activity in a particular sector.

OPEN MOVE

Italy

OpenMove

"Your app to travel freely"

openmove.com

Wikimedia uses all data available. Wikimedia is a global movement whose mission is to bring free educational content to the world and are best known as the organisation behind Wikipedia. What is less commonly known is their work on Wikidata, a free and Open Data-based knowledge base for everyone to access. The content of this database is distributed under a free license and can be linked to other Open Data sets. This database not only facilitates the automated dissemination of the knowledge gathered by the community, internally, it helps with the translation of Wikipedia articles into multiple languages. For external users, have such a vast amount of linked Open Data available enables them to visualise and analyse longitudinal or with a very broad scope.

WIKIMEDIA FOUNDATION

Belgium

Wikimedia

"The future of sustainability management"

wikimedia.org

The outcome of the Temperature Check survey largely confirms the findings above as can be seen in Figure 25. Respondents indicated in both surveys that the data category Government & public sector is the most re-used data category. Although not in the same order, the data categories Economy & finance, Population & society, Environment, Transport and Education, Culture & sport all appear in the top of most re-used data categories by both surveys. When looking at the private sector, most respondents of the Temperature Check indicated to mainly use data from the Government & public sector (41%), followed by the Economy & finance category and the Environment category (both 38%). On average, respondents from the private sector re-use a total of eight different categories when working with Open Data. Data sets from the Agriculture, fisheries, forestry & food categories are perceived as the least used (15%), followed by the Science & technology domain (18%). Most correlation between categories is found between the Government & public sector category (14 times), the Economy & finance category (10 times) and the Environment category (9 times).

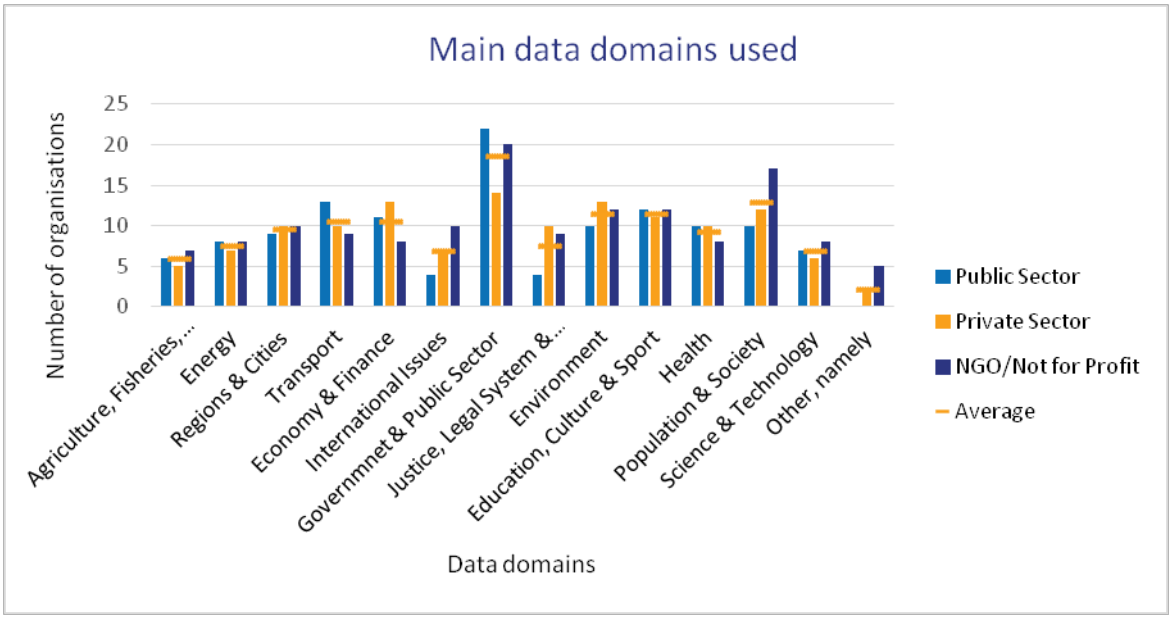


Figure 25: Main data categories used – IODC Temperature Check

Respondents from both the public sector and non-governmental organisations also indicated that the most re-used data sets are from the "Government & public sector" category. This category was positioned first by 71% of the respondents. The following categories perceived as the most re-used are Transport (42%) and Education (39%). The least re-used data sets come from the categories International issues and Justice (13% each). When considering the NGO sector, data sets from the Government & public sector were re-used most by 65% of the respondents, followed by the category Population (55%) and Environment and Education (39%). Again, the different organisations use multiple categories simultaneously. In the NGO sector, a strong correlation seems to exist between the category Government & public sector (20 times) and Population & society (15 times), whereas the public sector respondents perceive a stronger correlation between the categories Government & public sector (22 times) and Transport, Economy & finance or Education, culture & sport (all 10 times).

3.2 Where does the Data come from?

Companies are using data from a wide range of data categories, but where do they the find the data in the first place? As shown in Figure 26, organisations collect data from the whole of Europe, candidate countries and the United States. The United Kingdom and Spain seem to be the most popular countries from which most respondents collect Open Data to work with. An explanation for this strong representation is that organisations from both countries are overrepresented in this sample and the number of city level businesses. This in turn might be caused by the relative high levels of Open Data Maturity of the United Kingdom and Spain. Both countries can be categorised as trendsetters, offering a fertile soil for organisations to work with their Open Data[15].

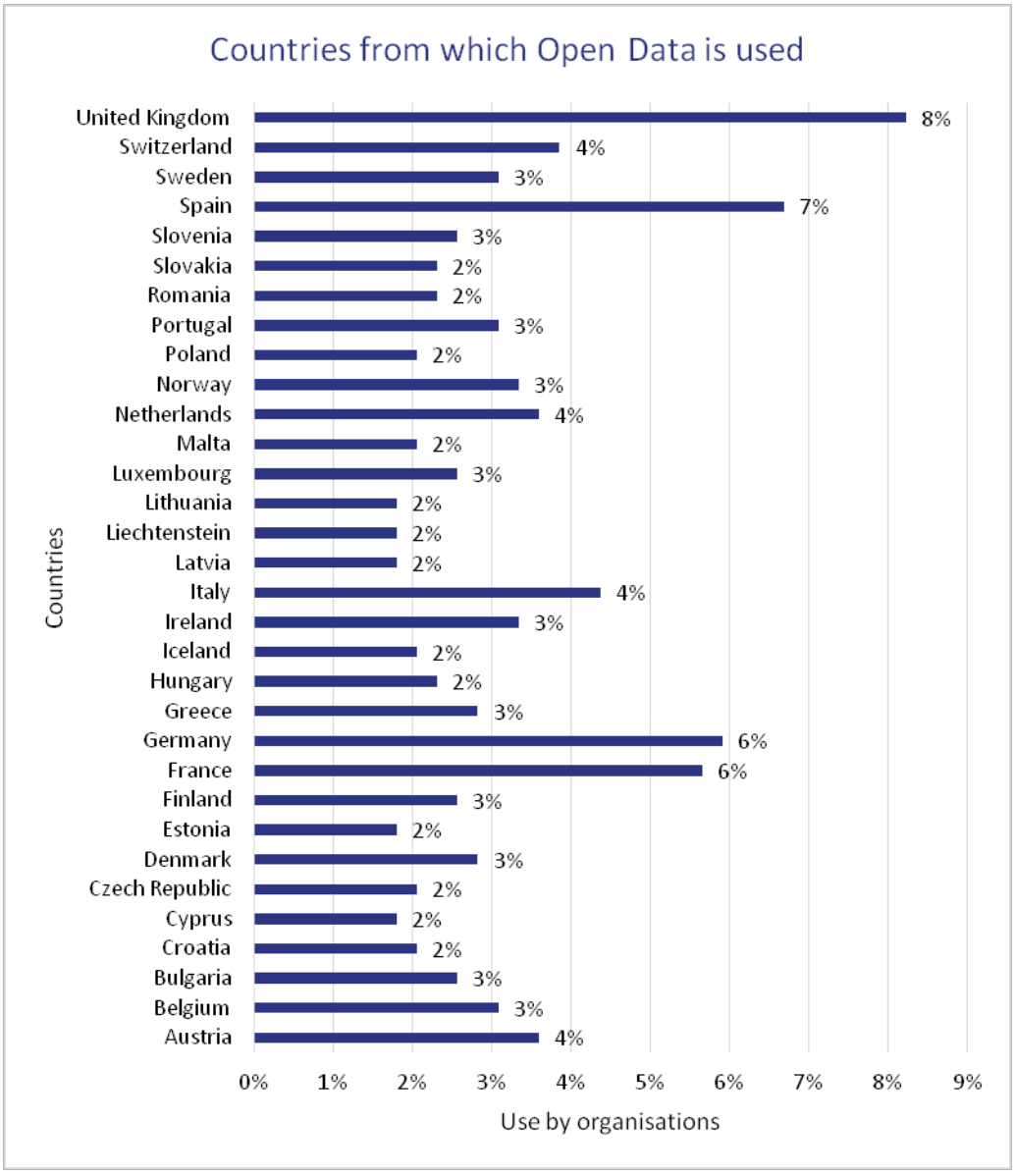


Figure 26: Countries from which Open Data is used (in percentages)

A strong variation in the responses can be identified, a variety which can be influenced by the type of business models the respondents have. For example the organisations that use Open Data to provide location specific services, such as **WiseTown**, will have a narrower geographical scope in terms of Open

[15] European Data Portal, 2016, Open Data Maturity in Europe

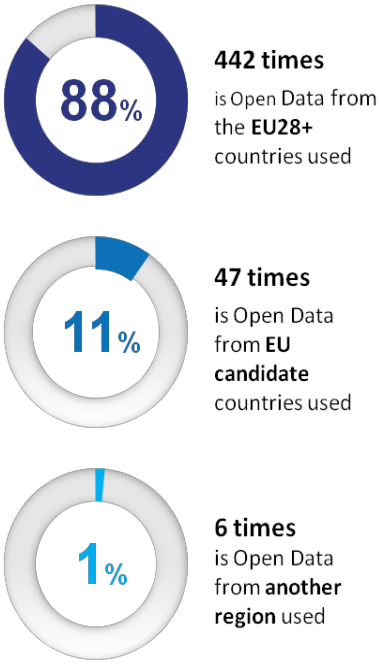


Figure 27: Regions Open Data is used from (in percentages)

Data compared to organisations that thematically collect Open Data such as **Infoempresa**. The business models therefore influence the scope of the required data. Taking this into account, the correlation between the location of the respondent and origin of its used data is assessed. When Spain and the United Kingdom are taken as examples to show this correlation[16], a conclusion can be drawn that over 80% of the respondents use Open Data from the country of establishment. This claim can be supported when looking at other countries represented in this survey by at least four companies per country: France, Germany, Italy and the Netherlands. The four French companies in this survey all use Open Data from France, while six out of eight German companies use data from Germany. In Italy, also six out of eight Italian companies use data from Italy. In the Netherlands, four out of five Dutch companies have indicated to use data from the Netherlands.

Open Data is not bound to national borders but is a global development, meaning our respondents can collect data from all regions in the world. Taking into account the strong representation of European organisations, it is no surprise that most respondents (88%) indicated to use data from the European Union region. What is more surprising is that, as can be seen in Figure 28, all regions are mentioned by the respondents. A likely explanation of this broad coverage comes from the fact that 38% of the organisations follow the business model of an Aggregator of Open Data. Another explanation lies in the fact that data from another geography can serve as a source of inspiration and lead a given company to compile similar datasets in their own region.

Companies re-using Open Data do not solely rely on the Open Data. Most often, businesses combine it with private data they hold directly or that is held by their client. This can for instance be seen in the work of BBVA Data & Analytics, Kenedict Innovation Analytics and Mozaika. BBVA Data & Analytics helps associated companies obtain insights in economic trends or the behavior of their clients by analysing and combining the data held by these companies together with available Open Data. Kenedict Innovation Analytics uses specific (private) data sources to conduct their analyses on the development of the market environment of their clients. These analyses are then further strengthened with findings from Open Data sources. Open Data provides them with additional insights and more context, enabling Kenedict to provide a more solid advice to their clients. Lastly, Mozaika combines and links data from other sources to Open Data. Again, it was stressed that linking Open Data to other (including privately held) data sources enables companies like Mozaika to get a clear view on the bigger picture. This enables them to conduct a more comprehensive analysis, leading to better informed decisions.

[16] The two countries which together represent 33% of respondents and collected data

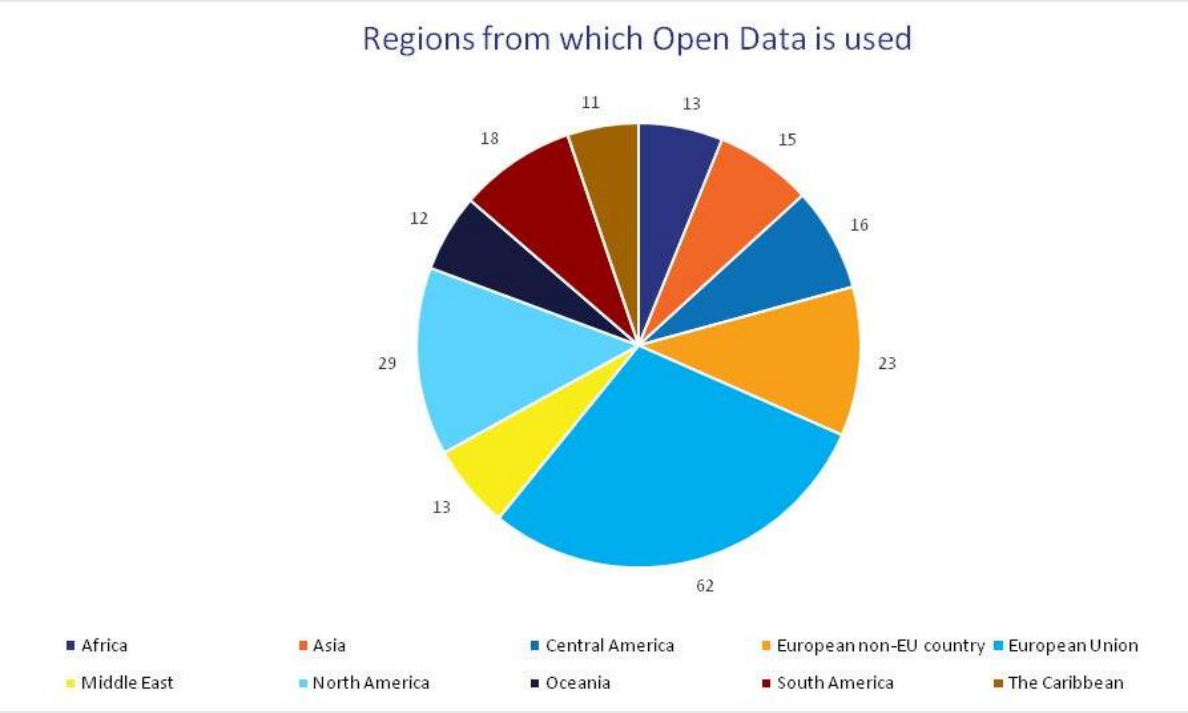


Figure 28: Regions from which Open Data is used (in absolute numbers)

3.3 The access to Open Data

After having looked at the geographical spread of the data being used, it is relevant to understand what types of platforms are used to access Open Data. National Open Data Portals are the most used platforms to access Open Data. However, despite the growth in data portals, the second most used source to collect data are the public administrations themselves. This could indicate a preference for the Open Data users to get their data from a platform close to the source where the data is produced. Or alternatively, that the discoverability of data on data portals remains a challenge. The third most used sources are regional or local data portals. This can be related to the fact that a number of companies offer geo-located services that may require specific data that can only be accessed via these portals. Many European cities have been launching their own data initiatives and offer direct access to their Open Data of which more information is presented in the report "Open Data in Cities[17]". Interestingly enough and perhaps correlated to the fact that all companies use data from at least two countries, the European Data Portal is the fourth most used. It represents a source for one quarter of the respondents. This is positive trend that will increase over time as the portal had only been live for 6 months when the survey was conducted. Figure 29 provides an overview of the different platforms used.

[17] European Data Portal, 2016, Open Data in Cities

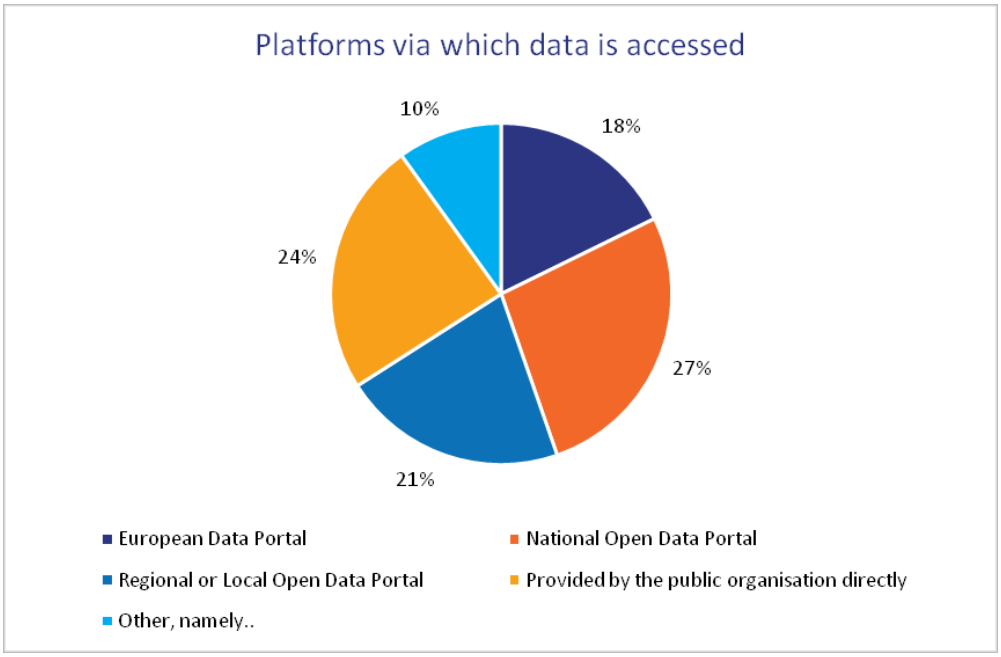


Figure 29: Platforms via which data is accessed (in percentages)

Fourteen organisations indicated to access the data by a different platform. These other platforms are primarily organised thematically instead of geographically, providing access to databases grouped on subjects such as oil, development aid, e-mobility facilities or judicial legislation. Typically these sources are OpenSensors, IATI Registry (on Development Aid), Open Charging Map and EURLex. Others indicate to either collect the data themselves, for example by the use of web crawlers, or by requesting particular data from the data holders directly.

A final insight is that on average, re-users use two different platforms (146 entries by 67 respondents) to access the Open Data. More specifically, 34% of the respondents indicated to re-use one platform to access Open Data, 28% indicated to use two platforms and 22% of the respondents uses three platforms. Furthermore, 52% of the organisations found it difficult to gain access to the Open Data they need, while 25% of the organisations indicated it to be easy. The use of more than one platform could be explained by difficulties experienced concerning the heterogeneity or variation in the Open Data, or low quality in terms of content. Another explanation could be the poor availability of certain datasets as the Open Data in certain fields can be limited or not consistently updated.

3.4 The benefits of working with Open Data

Working with Open Data brings multiple and synergetic benefits to organisations. The overall benefit of working with Open Data as mentioned by the organisations is the benefit Open Data brings to society in general. Open Data enhances government transparency and accountability as well as participation from civil society. Open Data can help bring together government and citizens which leads to a global benefit for society at large. For 16 interviewed companies this was seen as an additional benefit of working with Open Data.

At the company level, 12 out of 33 interviewed organisations explicitly mentioned Open Data to be the core of their business model. A prerequisite to work with Open Data is the availability of data and richness of information found in the data. This includes opportunities to use Linked Open Data[1] for machine learning and analysis as well as having information available from a certain scope or subject. As re-users of Open Data, the interviewed organisations identified the following five main benefits for companies working with Open Data. These benefits are visualised in the figure below.

- **Innovation** made possible by using Open Data. This was primarily the case in which the data used was previously closed and has now been used to develop entirely new business ventures or applications.
- **Reduced costs** and increased efficiency, by increasing data sharing practices across administrations, applying one time data provision principles and for example sharing projects with a public body that uses refined data based on its own Open Data.
- **Data harmonisation** done by companies to help public administrations to work on mapping the data they possess and determine how to publish it. This fosters data quality and harmonisation of metadata, use of controlled vocabularies, etc. leading to data that is discoverable and easier to use, in turn encouraging more data re-users to develop new products.
- **Enhanced business models** by using Open Data mainly relates to providing good quality and tailored made data as well as insights to businesses and clients in general. Open Data can also help save time as Open Data can be used to reduce the administrative burden of an organisation or eliminate time spent on requesting data.
- **Increased reliability** of the company since the Open Data used is official information from a public administration which creates trust and reliability to those making use of the services provided by the company. Examples are applications that make use of real-time arrival or departure data of public transportation services or applications that show where to find available parking spots or bike sharing stations in a city.

[1] Open Data by which the metadata can be connected and enriched, enabling to make links between different types of related resources.

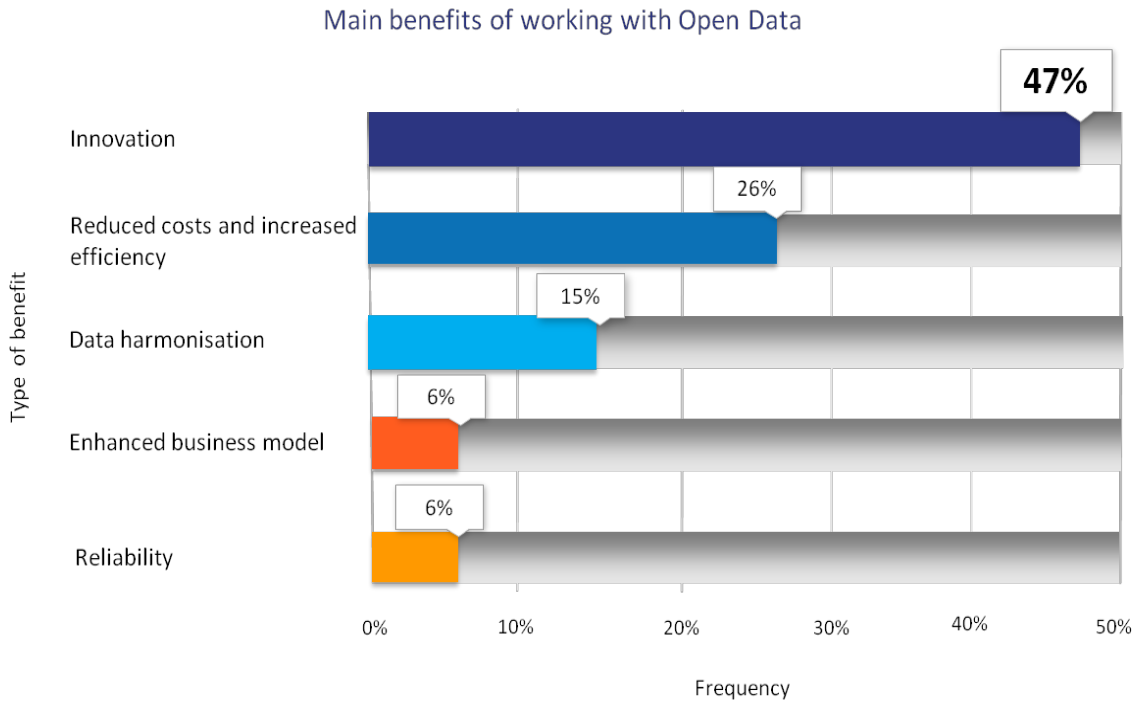


Figure 30: Main benefits of working with Open Data (in percentages)

The outcome of the Temperature Check confirms this view with a majority of respondents identifying the societal value as one of the main benefits when using Open Data. Especially in the public and NGO sector this was most apparent. In the public sector the societal value had been identified 22 out of 31 times as one of the main benefits, followed by creating new products (15 times). For the NGO sector, the difference was more significant, going from 23 out of 31 respondents indicating the societal value as one of the main benefits, to the scope of available information (10 times) and to creating new products (9 times). The private sector shows a more equal spread amongst the benefits, having 21 respondents indicating the societal value to be one of the main benefits, closely followed by the scope of available information and creating new products (both 19 times). The results of the temperature check concerning the main benefits are shown in Figure 31.

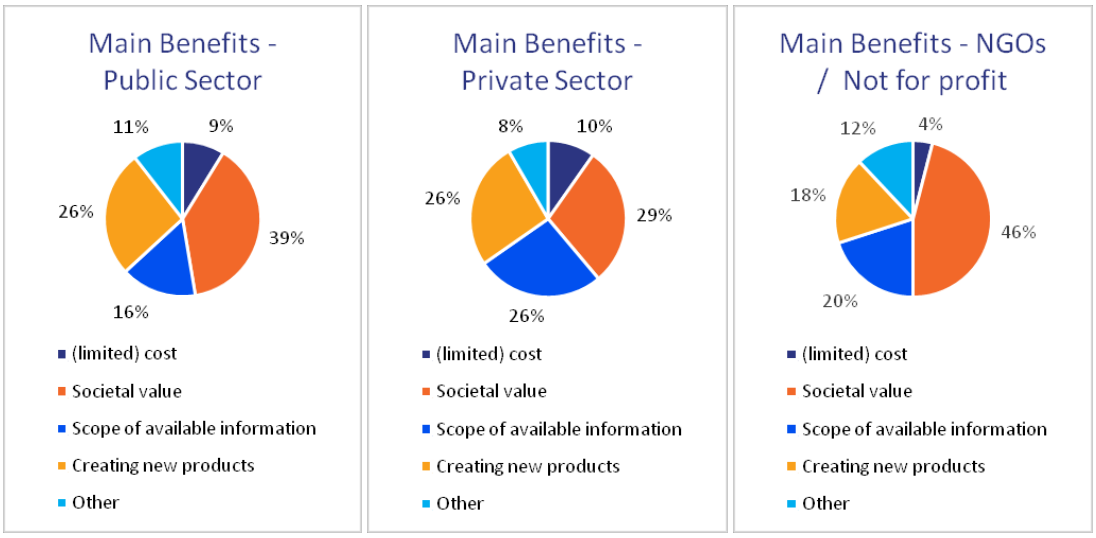


Figure 31: Main Benefits - Public Sector - Private Sector - NGOs/Not for profit – IODC Temperature Check (in percentages)

For the British company **Open Opps** the new possibilities of Open Data are the rationale of their existence. They bring this benefit into practice by using Open Data to centralise information on public tenders, offering one central point of access to companies looking for government contracting opportunities. With the available Open Data, they are able to collect and aggregate information on tenders and opportunities from 71 countries around the world. For them to offer this innovative service would not have been possible without having this information provided to them as Open Data.

United Kingdom

Open Opps

"Never miss a business opportunity"

openopps.com

The organisations were also given the opportunity to present the most interesting impact their organisation has established. The most frequently mentioned impact is that the respondents make information accessible which in turn enhances transparency, efficiency or awareness. This ranges from creating and publishing documentation from an entire region to becoming a reliable source for analysts to check their work. This is particularly valid in the case of aggregators. Others described the impact of their work from a societal point of view, such as the identification of potential public savings, contributing to the body of public knowledge or helping track and prevent illnesses from spreading.

For many of the companies interviewed, they see Open Data as being a core component of their activity. Open Data has been woven into their core services and/or products and is one of the key resources that has enabled them to start their business. But how do business strive financially in a context where their core resource is open and free to use by all and for any purpose? This will be explored in the next chapter.

EUROPEAN DATA PORTAL

Top 3 most used Open Data domains

27.3% Statistical

25.8% Geospatial

19.5% Companies

Most often combined Open Data categories

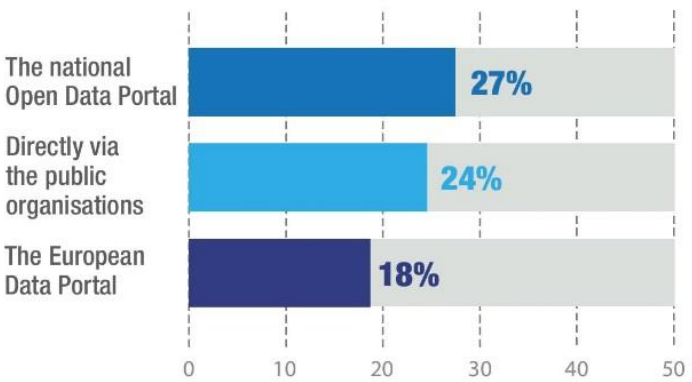
Other Open Data categories used

In clockwise order:

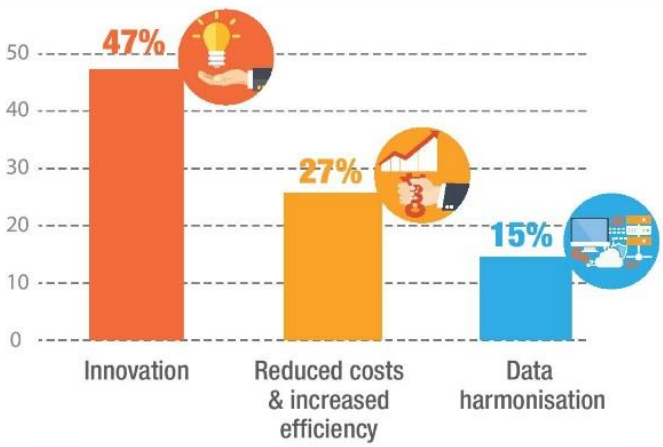
- Agriculture, fisheries, Forestry & Foods
- Energy
- Economy & Finance
- International Issues
- Government & Public Sector
- Justice, Legal System & Public Safety
- Education, Culture & Sport
- Health
- Science & Technology

Top 3 platforms to access Open Data

Provided by



Main benefits of working with Open Data



4. What business is there with Open Data?

Chapter 2 underlined that the benefits of Open Data are numerous, and the potential value of Open Data is significant. Chapter 3 highlighted the type of data that is re-used by the various organisations assessed. To gain further insight into the re-use of Open Data by public and private sector organisations, this chapter illustrates how the potential value of Open Data is exploited by these organisations in particular. Combining insights from both the survey, the interviews and the IODC temperature check, this chapter describes what re-users actually do with Open Data and how they materialise its value. This is further illustrated with success stories from a number of organisations. This chapter explores the business brought by Open Data, before exploring the different business models developed by organisations to generate revenue from Open Data. Finally, the assessment concludes on the capacity in terms of size and manpower organisations have to derive the value of Open Data.

4.1 What business does Open Data bring?

After having discovered what data businesses use and where they find it, it is important to understand what businesses do with the data and where they make profit.

4.1.1 Selling services as the most important source of revenue

To understand by which means organisations generate revenue by (re-)using Open Data, they were asked to indicate what their significant source of revenue is. As can be seen in Figure 32, the majority (21%) obtains their revenue by selling services or products or a combination of both. How these services are translated into revenue, for example by which distribution channel or payment system, will explored in section The role of Open Data in business models 4.2 where the business models behind the organisations will be explored.

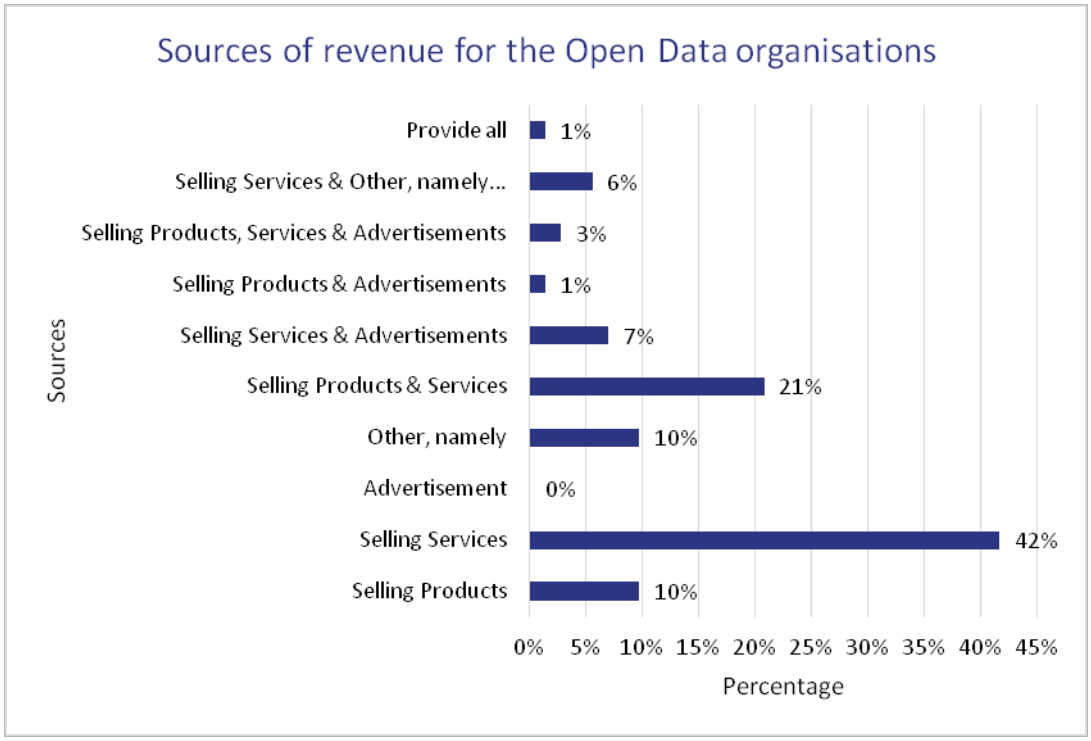


Figure 32: Sources of revenue for the Open Data organisations (in percentages)

In addition to the sources of revenue indicated in Figure 32, the organisations provided descriptions of their revenue streams which revealed additional insights. By clustering the descriptions into themes,

a trend can be identified. The following themes were identified: 1) Consulting, 2) Software as a service, product or subscription, 3) Access via API, 4) Donations, 5) publicly funded, 6) Selling services, 7) Selling products, 8) Other. 34% of the organisations named software as their most significant source of revenue (Figure 33). Software is often commercialised as a service (SaaS) but also by subscriptions or by licensing the use of it.

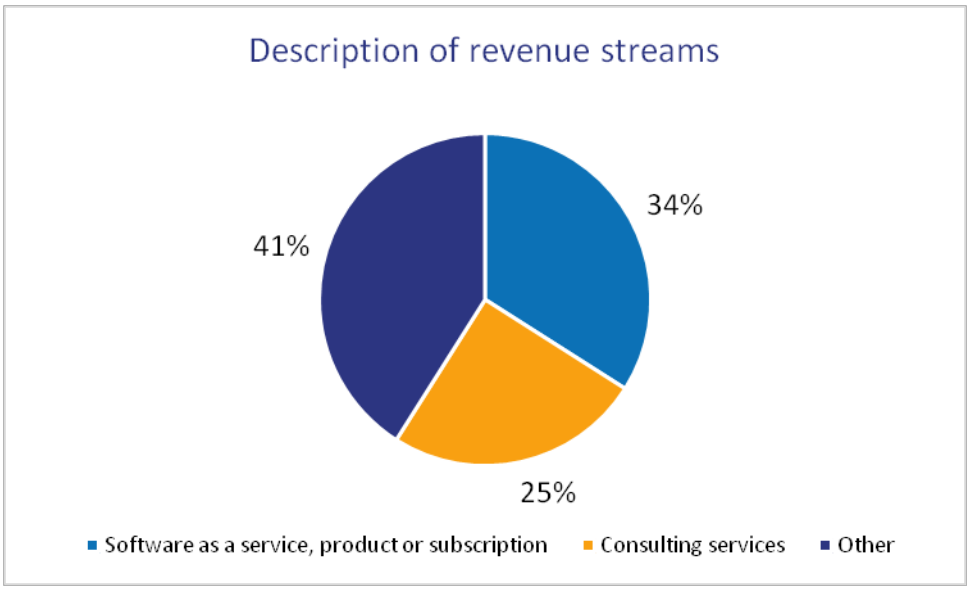


Figure 33: Clustering of qualitative descriptions of revenue streams

Regarding the second most significant source of income, 25% indicated to create revenue by offering consulting services. 41% of the respondents gave either highly specific sources of revenue (e.g. interest rate margin), a description of the sector of their clientele (e.g. e-mobility) or their product in detail and were therefore classified as 'other'. The different sources of revenue – software, consulting services and 'other' can be illustrated by the following examples.

An example of an organisation selling both services and products based on Open Data is the **Connemara Programme**. The Connemara Programme offers a variety of Open Data based services and products in the Connemara region of Ireland. Using digital information about the economic, social, physical, cultural and natural environment, Connemara businesses are supplied with ready to use tools to enhance their presence on the internet. To do this, the company has collected and categorised information about the whole 2000 km2 region. With this, they offer map based online directories, data sets and support services to businesses, community organisations, individuals and policy makers.

Ireland

Connemara Programme

"Making Connemara visible to the world"

connemaraprogramme.com

Germany

GraphDefined

"Information and Communication solutions for a digitized open world"

graphdefined.de

The example of e-mobility as a description of a very specific sector-focus is best illustrated by **GraphDefined**. This is a Germany company that consults on and creates software for all participants in the Open Data Value Chain. As a strong believer in Open Data, this company is working on the development of new business models by which companies are encouraged to publish and re-use information. For example, by developing applications which map the location and status of e-mobility charging stations in which not only the supplier of the electricity receives financial remuneration but also the supplier of the (Open) Data on the charging station in question. With incentives like these, it also becomes interesting for the data holder to share his information and leverage the benefit of aggregated data.

An example of a company selling services based on Open Data is **Geomanagement**. This Greek company offers various services based on the analysis of geospatial data, for example from the Copernicus Project (http://www.copernicus.eu/). With these analyses the company helps clients make their work easier or to find solutions for specific problems. By combining data analytics with available Open Data, the company helps both public and private sector clients. For example, one analysis was conducted to demonstrate how Greek farmers were affected by forest fires or floods. With this information, they could be exempted from certain taxes or made eligible for compensations.

Greece

Geomanagement

"IT Open Source Solutions"

levadia.gr

The NACE-sector[18] of the organisations is directly related to what the organisations sell and the domains they sell their services or products in. Over half of the respondents are active in the Information and Communication sector or identify themselves of providing professional, scientific and technical activities. As can be seen in Figure 34, all the other sectors are significantly less represented.

[18] The NACE Code (Nomenclature statistique des activités économiques dans la Communauté européenne) is the industry standard classification system used in the European Union. The full list of possible NACE sectors for the respondents to select is displayed in Annex II.

Although the benefits of Open Data are apparent in multiple sectors, the organisations working with this resource are still predominantly clustered in the ICT sphere. An example of an organisation which serves other sectors while being active in another is **CropDiagnosis**, which is presented on the next page and develops applications for the agricultural domain.

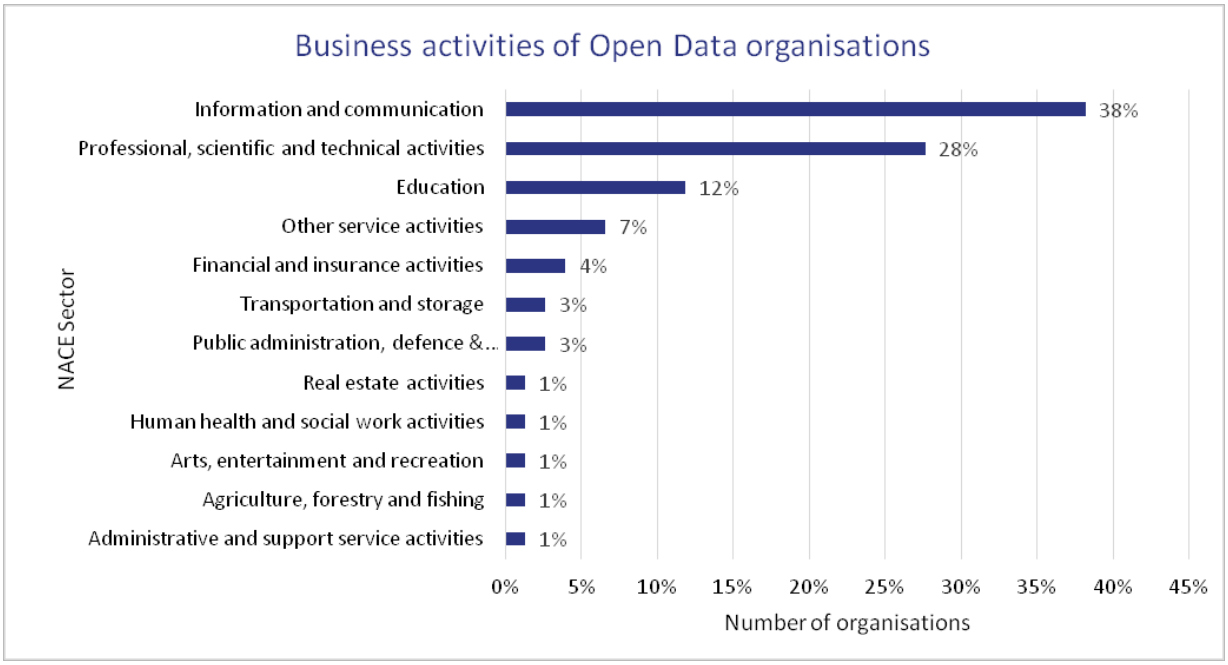


Figure 34: Business activities of Open Data organisations (in percentages)

The clustering of organisations in the ICT sector directly relates to the fact a large proportion of the respondents sell software as a service; as well as consulting in the field of data analytics. However a number of organisations have a sectorial focus, as their core business and focus is on safety and security such as **Netage** and **Canard Drones**, education and public service innovation such as **eGovlab** or on financial services, such as **BBVA**. Despite this sectorial focus, all organisations made use of ICT and ICT related skills to retrieve and process the data. Chapter 5 provides more insight in the skills needed to do this.

Greece

CropDiagnosis

"The easiest way to keep your plants healthy"

cropdiagnosis.com

CropDiagnosis is a good example of a company providing a technical 'Information and communication' service for a sector completely different. This Greek application, a spinoff from Ergobyte, helps farmers diagnose diseases, identify threats and achieve optimal crop management and treatment of the harmful contaminations. By asking the farmer a series of specifying questions, the application runs multiple queries on an extensive database composed of multiple Open Data sources. With these queries, the crop's details are determined and the threat is characterised in order to make the diagnosis and provide the best recommendation to mitigate the issue.

The Spanish company **Canard Drones** takes it up in the air to provide their services. With the use of a particular type of Open Data published by nearly every airport in the world, they help airports calibrate their aeronautical navigation aids to pilots need to land safely. What is special about this services is that Canard Drones does not calibrate these aids in person or manually but that they use unmanned aerial vehicles for this purpose. By using drones and by planning the operations meticulously, for using all types of Open Data available, this is much more efficient than traditional ways of calibrating the aids. Not only does this means lower costs, using the electrical drones also means that it is a 0% emissions procedure, not only saving costs but also contributing to the sustainability of the environment.

CANARD
Spain

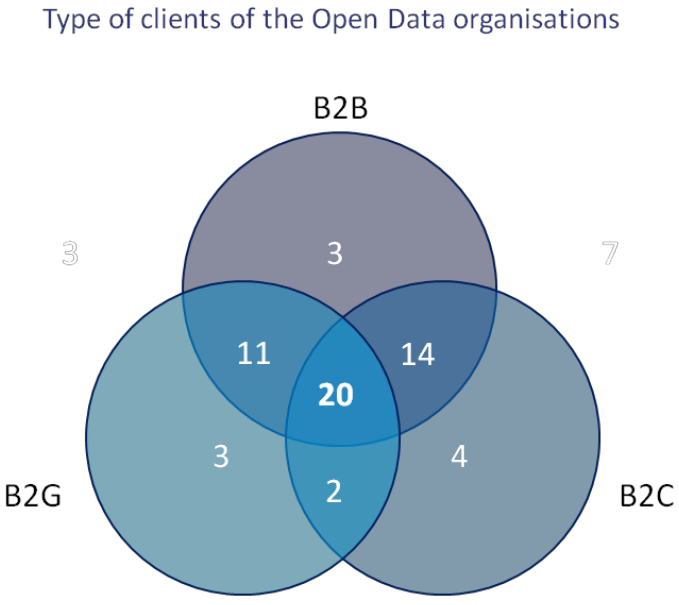
Canard Drones

"Calibration of aeronautical Navigation Aids using drones"

canarddrones.com

4.1.2 Who are the clients?

The businesses surveyed were also asked to define to whom they deliver their goods or services, e.g. to other businesses (B2B), consumers (B2C), governmental organisations (B2G) or a combination of the above (Figure 35). Regarding this question, the versatility of Open Data was not manifest. A little under one third serve all client segments while more than half serve two or more client segments.



Type of client per organisation, absolute numbers
N= 67 (10 organisations indicated 'other')

Figure 35: Client segments

This finding could indicate that the services offered by the organisations are not limited to certain domains nor groups. An interesting company in this respect was **SmartAppCity**, which is presented in section 6.2 as this company is explicitly serving both the governmental and business sector with their application. They do so by selling their service as a private-public partnership in which the costs are shared between the two.

4.1.3 Large discrepancies remain when investigating share of Open Data in turnover

The financial performance of the organisations was measured in terms of turnover. This was assessed for the last available year and the average turnover of the respondents over the last 5 years. As can be seen in Figure 36[19], both the figures of the most recent and average turnover show a trend of annual turnovers under €49,999. This low figure can be partly explained by the youth of the organisations who took part in the survey, as half of them did not yet exist 5 years ago.[20]

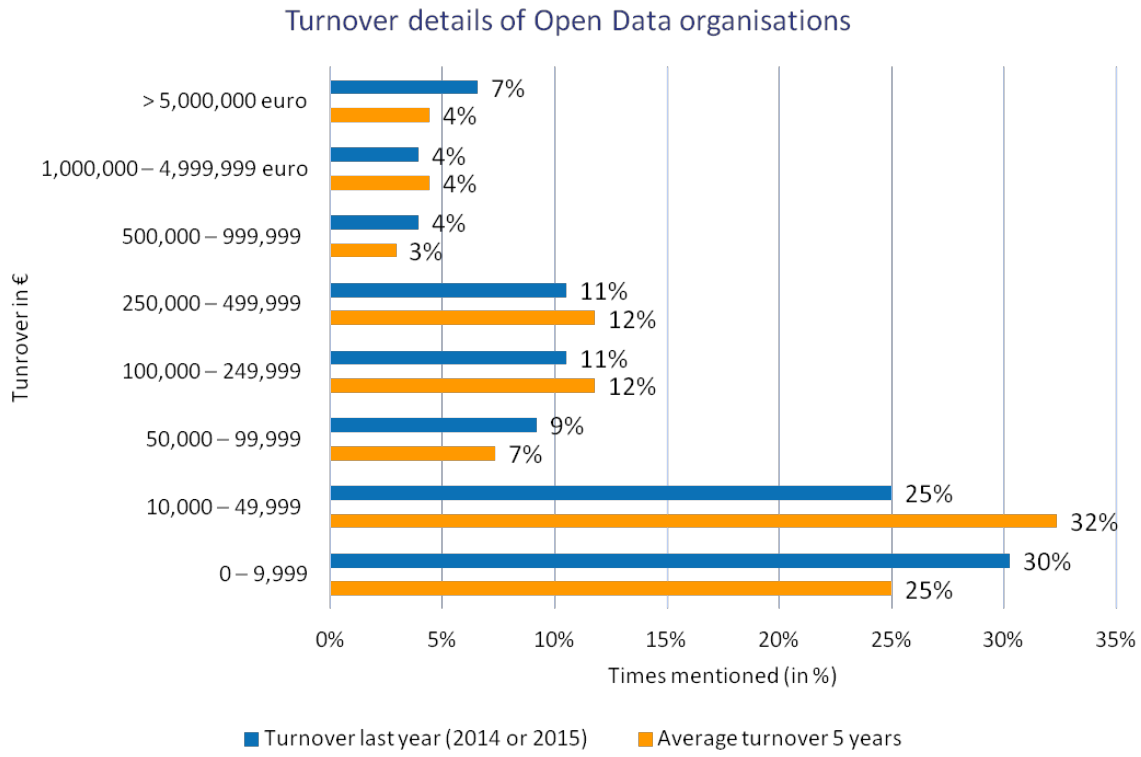


Figure 36: Turnover details of Open Data organisations (in percentages)

To obtain a clearer picture of the impact of Open Data, organisations were also asked to indicate how much of their turnover is related to the specific use of Open Data. To facilitate both the organisations whose businesses are completely based on Open Data and those who use it to improve existing business models, organisations provided indications for both the direct and indirect percentage of turnover related to the use of Open Data. As shown by Figure 37, for most organisations Open Data only has a marginal influence on their turnover. A possible explanation for this low figure is that most organisations surveyed primarily offer certain services (e.g. software or research) for which Open Data can be used such as research services or Internet of Things (IoT) solutions. This means that although they use Open Data, the turnover is primarily generated by the 'core' service and Open Data is only used as an extra.

[19] To prevent statistical errors, the output of respondents which only provided information on the most recent year are removed

[20] As can be seen in Figure 59, 39 respondents indicated to be founded in or later than 2012.

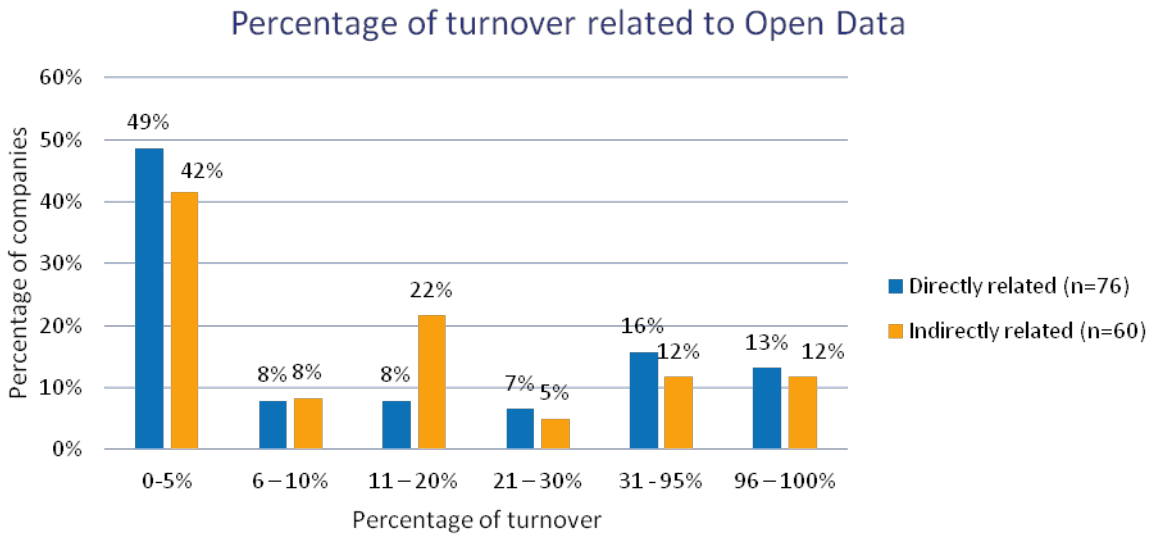


Figure 37: Percentage of turnover related to Open Data (in percentages)

However looking into the numbers in more detail; for 23 of the respondents, that is over 30% of the sample surveyed, Open Data does directly represent a substantial part of their turnover, ranging from 11% to 95%. Furthermore, a specific sample of 10 organisations did report that Open Data revenue could represent 96% to 100% of their turnover. Larger companies tend to diversify their revenue, thus, as mentioned above, may be seen as diluting the share of Open Data in the overall turnover.

Similar conclusions can be drawn when looking at the turnover which is indirectly related to Open Data. Here, 42% of the organisations reported that only a maximum of 5% of their turnover is indirectly linked to Open Data. But at the same time, out of the 60 organisations in this sample, 23 indicated that 11% to 95% of their turnover can be indirectly linked to Open Data.

Considering the relative youth of the organisations, it is likely that the figures are distorted by the fact that most organisations are in a start-up phase. During this phase, many organisations state to be heavily investing in product development with a relatively low turnover. But how do they see the future, and what is the role of Open Data in their growth projections? Figure 38: Forecasted annual growth in general and related to Open Data (in percentages)

Figure 38 displays how the respondents foresee the growth of their organisation in the near future, both related and unrelated to Open Data.

When looking at the forecasted growth, the respondents show diverse growth prospects. A quarter of the respondents is forecasting their general turnover to grow between 0-5%. But at the same time, 37% expect to see their turnover to grow by 61% or more in the coming years. For the cohort with the highest growth forecast (> 300%) half of them indicated that their growth related to Open Data would only be 0-5% while the others expected Open Data to play a bigger role of up to 100%. Although there are differences in forecasted annual growth between the general turnover and the turnover related to Open Data, this difference is not significant.

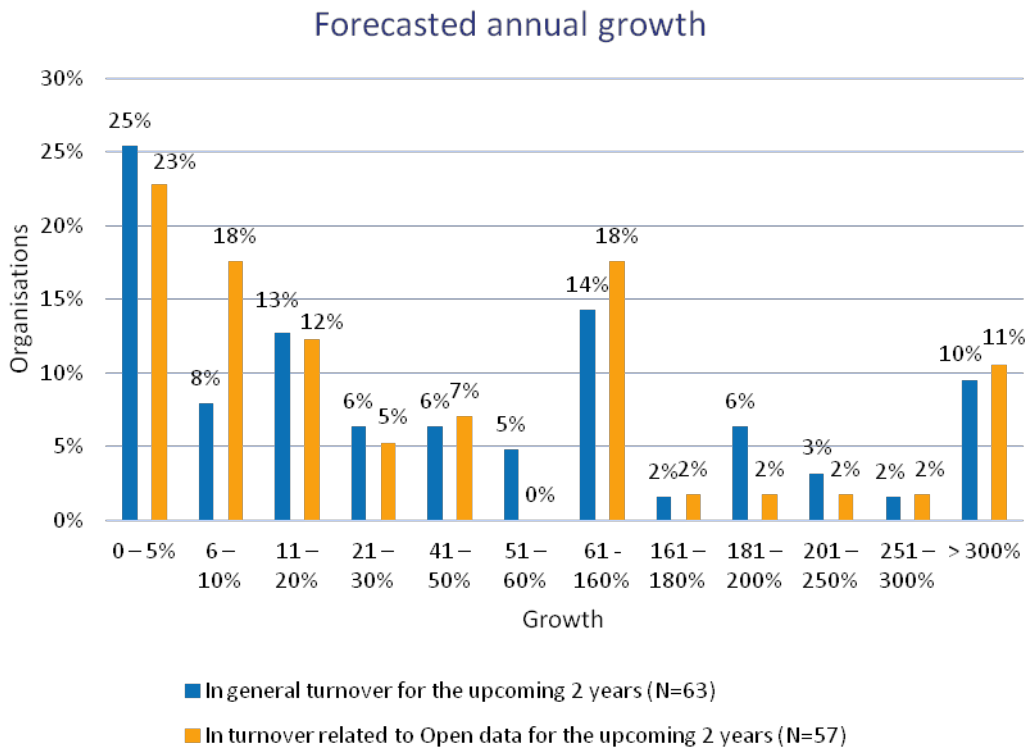


Figure 38: Forecasted annual growth in general and related to Open Data (in percentages)

Taking into account the average youth of the organisations, the forecasted growth was also examined by juxtaposing the age of the responding organisation. Neither by the examination of the forecasted growth rates per year nor vice versa, trends were identified. For example when looking at 2014, the most frequently mentioned founding year with 16 organisations, ten different growth rates are forecasted ranging from 0 – 5% to > 300%.

When the forecasts of the respondents in the two most frequently mentioned NACE sectors are examined, the most frequently mentioned sector – Information and Communications – displays a similar spread of growth expectations. Of companies active in the second most popular sector - Professional, scientific and technical activities – half of the organisations have a forecasted growth of 0-5%.

Comparing the growth forecasts against the Open Data percentage for their latest turnover shows that the share of Open Data is expected to remain stable in over two-third of the organisations surveyed. Interestingly enough, most of them did not systematically choose to de-couple the growth they expected in terms of general turnover from the growth directly related to Open Data. This is an encouraging finding as it points to the fact that a majority of the respondents see the share of Open Data in their turn over as a continuous driver for growth. Furthermore, as show in Figure 39, 25% of the organisations surveyed clearly expect Open Data to be a substantial driver of additional growth.

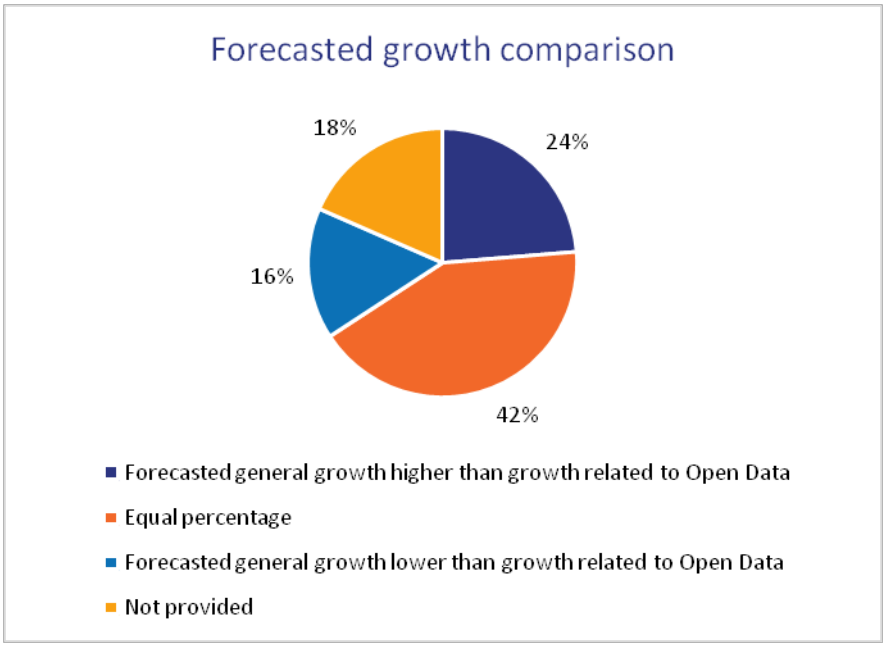


Figure 39: Forecasted general growth compared to the growth related to Open Data

When looking at the influence of Open Data on the turnover of 2015[21], this trend is already visible as ten organisations indicate a higher indirect percentage than direct percentage of Open Data on their turnover. A possible and partial explanation for this growth comes from organisations using Open Data for the development of proof of concepts that have only been partially commercialised. Once this phase is completed, it should be easier for them to firstly increase their turnover from the sales of full-fledged products and services and secondly expand their use of data to other product or service.

4.2 The role of Open Data in business models

Going more in depth on Open Data, the respondents were asked about the role Open Data plays in the business model of their organisation. To determine this, they were asked to indicate where Open Data adds value and which (Open Data) business model description fits best their organisation.

4.2.1 The emergence of data brokers

In order to identify which business model description fits them best, the respondents could choose out of the archetypes identified by several studies in the Data Value Chain[22] and displayed in Figure 9. These archetypes follow the steps of the Data Value Chain to identify at which moment the organisation is involved with the data, as explained in Figure 10 earlier in this report.

As can be seen in Figure 40, most of the organisations classify themselves as an aggregator (36%), working on the collection and aggregation of data. The remaining classify themselves almost evenly into one of the other archetypes, ranging from 14% to 17% per archetype. This is a surprising outcome as the study expected to find primarily organisations from the right side of the spectrum of the Open Data Value Chain, using Open Data as a resource for an application. These results clearly underline the intermediary role played by private sector companies in offering services as data brokers, confirming the emergence of an infomediary industry.

[21] Or for the year 2014 if 2015 was not yet available

[22] MEPSIR (2006), p. 46 and http://ec.europa.eu/information_society/newsroom/cf/dae/ document.cfm?doc_id=3488 http://www.worldbank.org/content/dam/Worldbank/document/Open-Data-for-Economic-Growth.pdf

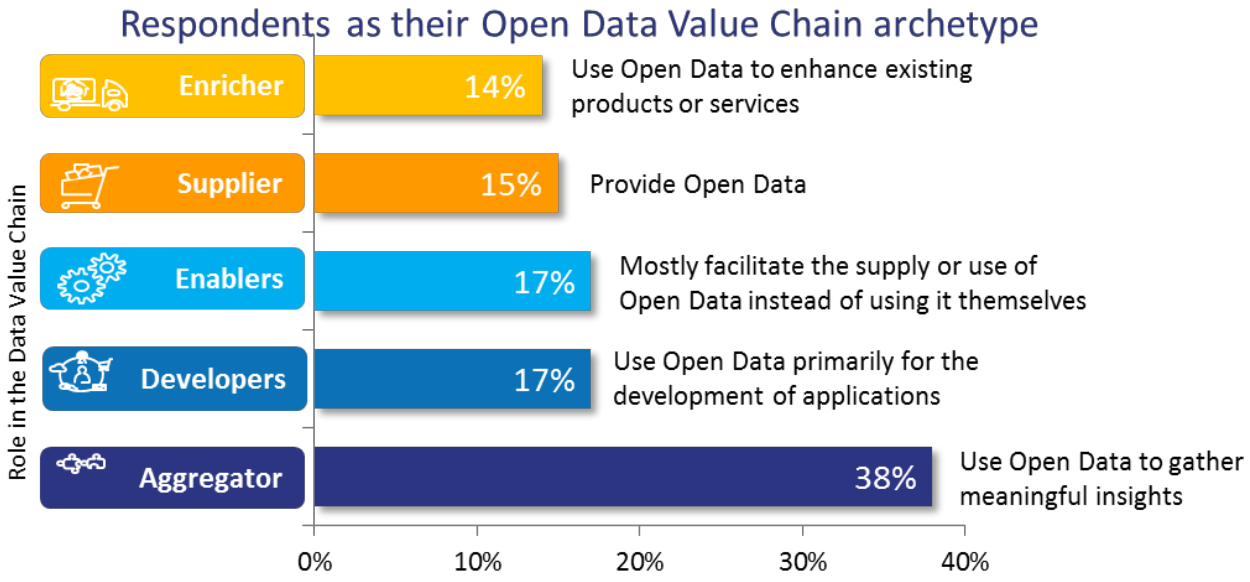


Figure 40: Respondents as archetypes of Open Data

Identifying oneself as an aggregator is supported by the outcome of the Temperature Check conducted at the very final phase of the study (Figure 41). When looking at the private sector, most Open Data users indicate to be a data aggregator (26% of the private sector organisations). This is followed by data suppliers (21%), data enablers (21%), data enrichers (19%) and least data developers (13%).

For the public sector, we found that most Open Data users identify themselves as data suppliers (36%), followed by data enablers (20%), data aggregators (18%), data enrichers (14%) and also data developers the least (11%). Data aggregators appear to be least active in the Non-Governmental Sector, having only 15% identifying themselves as such. In this type of organisation, most Open Data users identify themselves as data enablers (26%), followed by data suppliers, data developers and data enrichers (all 20%). This is shown in Figure 41 below.

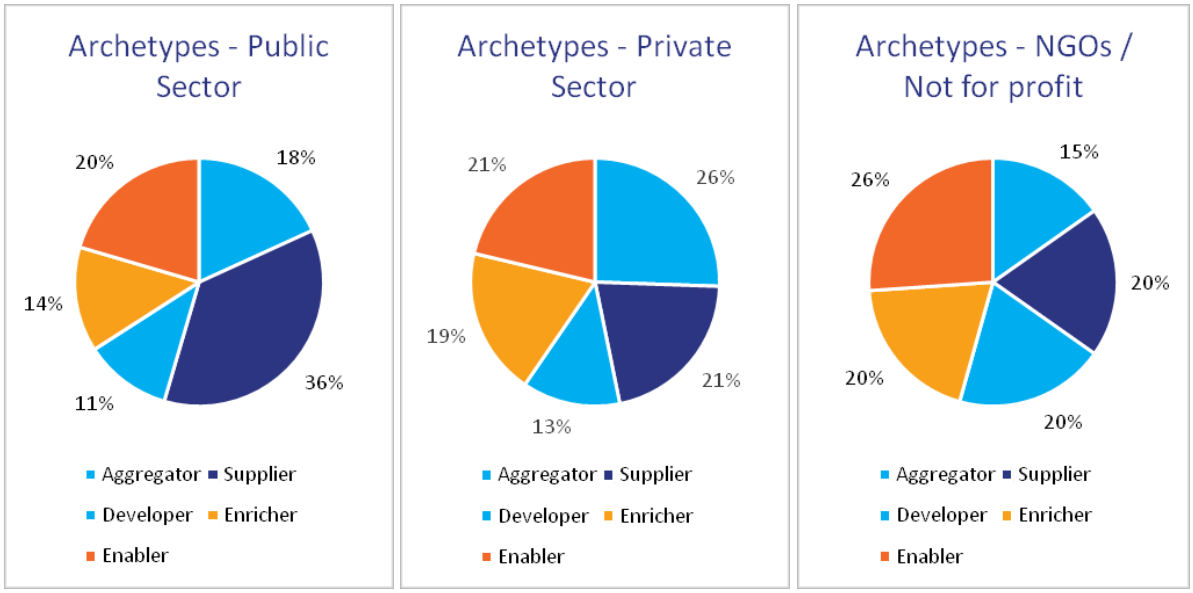
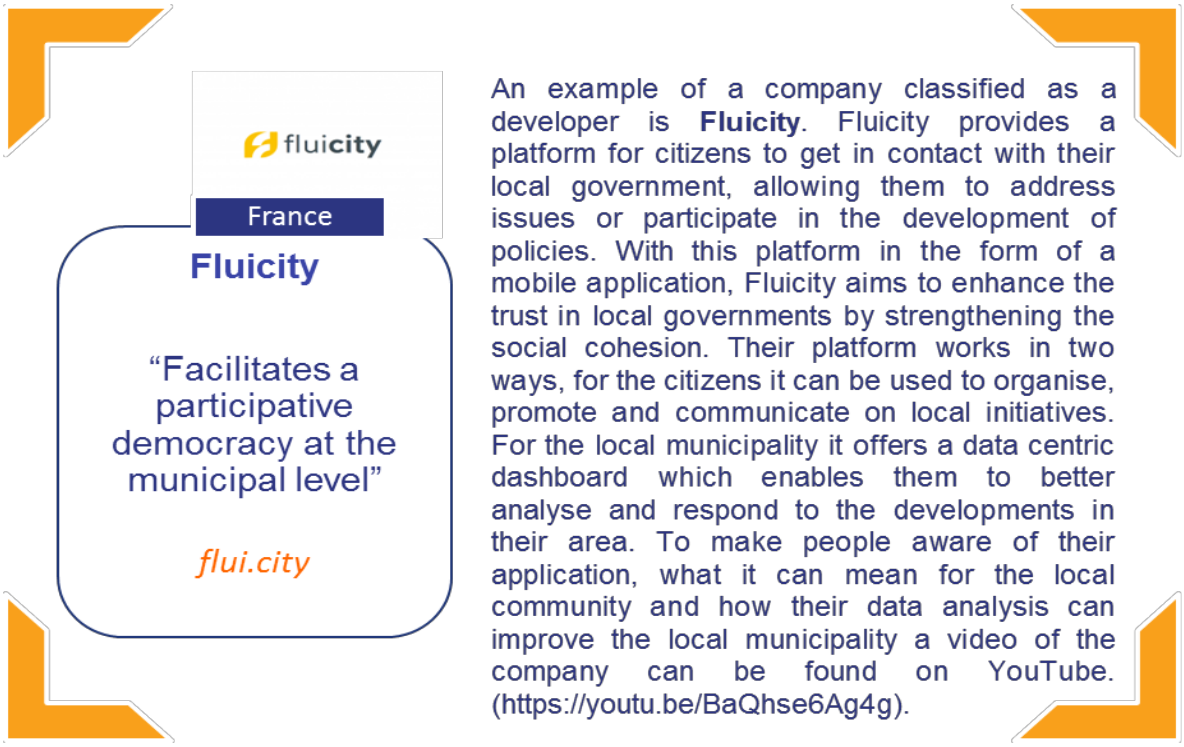


Figure 41: Archetypes - Private sector - Public Sector - NGOs/Not for for profit – IODC Temperature Check (in percentages)

This means that it is important to take into account the nature of the organisation when interpreting the results of Figure 40. While the results of private sector organisations included in the Temperature Check are roughly in line with the results of the survey, public sector organisations and NGO's reported different dominant archetypes. While re-using data, public sector organisations at the same time identify themselves as data suppliers. This was expected, as the public sector creates the data as part of their daily operation. As compared to private organisations, data aggregators are less common among both public sector organisations and NGOs. These type of organisations are more active in the 'supply-side' of the data value chain.



An example of a company classified as a developer is **Fluicity**. Fluicity provides a platform for citizens to get in contact with their local government, allowing them to address issues or participate in the development of policies. With this platform in the form of a mobile application, Fluicity aims to enhance the trust in local governments by strengthening the social cohesion. Their platform works in two ways, for the citizens it can be used to organise, promote and communicate on local initiatives. For the local municipality it offers a data centric dashboard which enables them to better analyse and respond to the developments in their area. To make people aware of their application, what it can mean for the local community and how their data analysis can improve the local municipality a video of the company can be found on YouTube. (https://youtu.be/BaQhse6Ag4g).

A company that can be classified as an aggregator is **openlaws** from Austria. As a big company, operating in more than one country, there are a multitude of compliancy requirements which have to be satisfied. With an ever growing body of case law, legislation and papers from legal scholars it can be hard to keep an overview of what is relevant for your company. By providing one central point of access to multiple Open Data based databases with legal information, openlaws helps legal professionals to search, organise and share the legal information he needs. With this function, they serve as a facilitator for practitioners to access and harness the potential of the available Open Data on legal subjects, for themselves or for the society.

4.2.2 Organisations are combining business models

There were seven options in the survey for respondents to indicate how they integrate Open Data in their business model. Organisations could depict their activity as 1) Enhancing products, 2) Enhancing services, 3) Process optimization, 4) Data as a service, 5) Information as a service, 6) Answers as a service, 7) Development of web or mobile applications; or several of these. Open Data can for instance be used to offer Data as a Service, to automate processes or for the development of web or mobile applications. The most recurrent business model around the use of Open Data is 'Enhancing services' (21%, Figure 42).

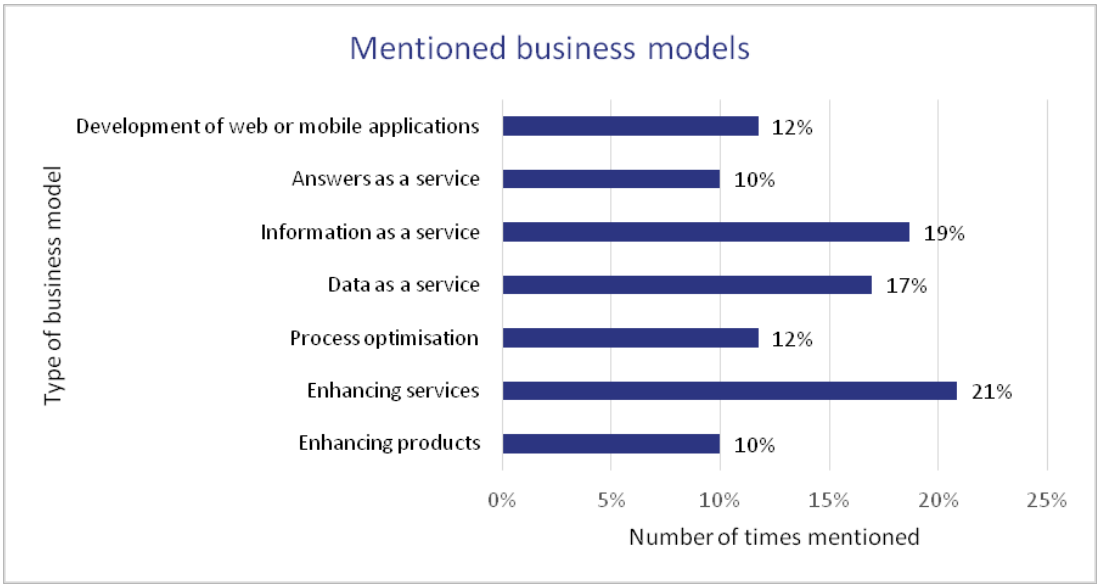
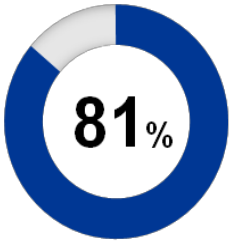
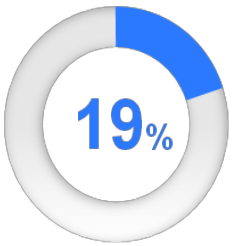


Figure 42: Mentioned business models (in percentages)

Mentioned slightly less, organisations use Open Data to provide Information as a service (19%) or to provide Data as a Service (17%). The other ways to create value using Open Data are mentioned significantly less frequent, between 10% to 12%. Organisations often combine these elements in their business model. 81% of the organisations indicated to use two or more of these business model elements.



81% of the organisations indicated to use two or more business models



19% of the organisations indicated to use only one particular business model

Figure 43: Organisations using one or multiple business models (in percentages)

As shown in Figure 43, by far most of the respondents use Open Data in multiple ways, using two or more business models. One of them even indicated to use Open Data in all possible ways. Taking into account the archetypes by which the respondents classified themselves, this outcome is counter intuitive. While the popularity of using Open Data to provide data or information as a service reflect the archetype of an Aggregator, this is not the case for 'Enhancing Services' (Figure 44). This approach to create added value using Open Data would normally be found within the archetype of Enrichers. Considering the popularity of using Open Data in multiple ways, this inconsistency could be explained if the enhancement of services is a secondary or lower priority for most of the respondents.

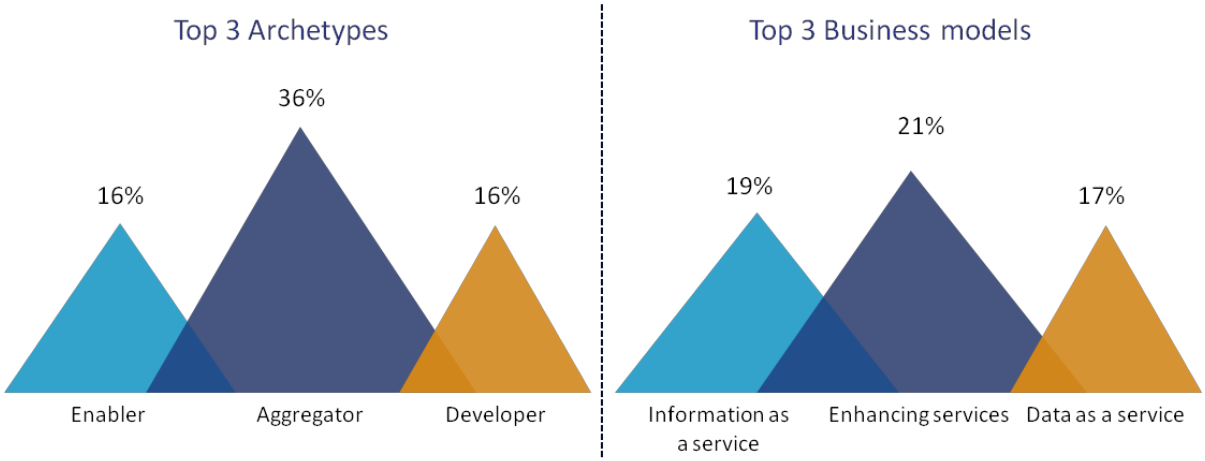


Figure 44: Comparison between archetypes and business models (in percentages)

For many of the organisations under consideration, services are offered in a freemium business model. With this model, a selection of the services is available to all without costs after which the full suite of services is only available to paying customers. This way potential clients can discover the value of the services offered.

An example of a company that indicated different business models – among others, 'Enhancing Services' - are applicable, is **Appchallenge**. This company from the United Kingdom is helping companies leverage the power of Open Data by bringing them in contact with the crowd. They do so in a structured manner, helping their client understanding the potential of their data and disclosing it by an appropriate access infrastructure such as an API. To transform the potential of the client's data into a tangible application or service, AppChallenge then harnesses the power of the crowd by organising app competitions. These competitions are run within their community of app developers, during which the participants can win an award for the most innovative or suitable application. To share their knowledge on Open Data, their competitions and the innovative solutions which have evolved from them AppChallenge also runs a YouTube video channel (www.youtube.com/appchallenge).

AppChallenge

United Kingdom

AppChallenge

"Runs competitions to help clients get the best Open Data application"

appchallenge.net

4.2.3 Adding value with the use of Open Data

The results of the interviews are very much in line with the findings of the survey. The interviews offered more granularity in four specific re-uses of Open Data:

1. Organisations who use Open Data for the enhancement of internal processes
2. Organisations who facilitate access to and services on (aggregated) Open Data for others
3. Organisations who offer (data analytical) products and services based on Open Data
4. Organisations who do not add financial value but create societal value

The first way by means of which Open Data can add value is by using the information stored within it to **enhance the (internal processes)** of a given organisation – in line with the category 'Enrichers' in Chapter 4.2.1. By analysing Open Data, a number of internal processes can be changed or improved to make them more efficient or suitable to the task. BBVA Data & Analytics is an example of a company that uses Open Data to enhance internal processes. They are a separate entity within the bigger financial group of BBVA, and focus on exploiting technology and analytics to optimise their processes. To help fellow group companies in the development of new or more services, BBVA Data & Analytics analyses the data of these companies in combination with available Open Data. Based on these analyses, associated companies are offered insights on economic trends or the behaviour of their clients. With the large amounts of data they gather by the provision of their services, the insights which this can reveal can also be of interest for other parties. To provide selected partners from the public or private sector access to this potentially valuable information, BBVA therefore shares raw, aggregated and anonymised data through an Open Data Platform.[23] This example shows that while not directly commercialising Open Data into a new product, it also creates value when used for the enhancement of existing products or services.

The second recurring business model is **facilitating access** to the information stored as Open Data and/or provide services to work with Open Data such as data curation. This form of value creation with Open Data has similarities with the work of the aggregator or supplier archetype of the Open Data Value chain. Aggregators make it possible or easier for other parties to use and analyse data, for which their clients are willing to pay. Where Wikimedia wants to provide access to nearly all subjects, the Connemara Programme has a strong focus on collecting and providing access to Open Data on a given geographical region.

A good example of a company providing access to Open Data on numerous subjects is the English company **Unigraph**. This company provides access to unified and interconnected information from multiple Open Data sources by the knowledge graph they have created. For their clients the value of using the service is the convenience of having the data they need in a cleaned, organised and real-time fashion. This saves them time and costs of going through the process of collecting and cleaning the data themselves. The body of knowledge stored in their graph is available to their clients over a paid API but also as a free Open Data 'dump' for others to install it on a different infrastructure for people to use it.

[23] https://www.bbvaapimarket.com/

For most of the organisations, it is not solely the information collected which adds value but also the services that can be provided based on the aggregated information. By having aggregated information on a given subject, services can be offered which visualise or provide insights on the data subjects. The client in turn has data they can combine with their own business insights and use for decision making. An example of how this can lead to value-add is shown by the Spanish company Infoempresa.com, part of the cloud and mobile technology provider Telecoming. While for this parent company working with Open Data is not their primary business occupation, it is noteworthy to mention the award they received for being the best job-creating Spanish SME, underlining the potential of the company and Open Data.

Infoempresa.com, the web service working with Open Data brings together information on Spanish and European companies from a multitude of Open Data sources, standardises the format of this data and processes it to provide the end user a searchable database and valuable insights for their business decisions. In this database, users cannot only find information on the financial, legal or other relevant performances of the companies, it also maps the relations between the administrators of these companies. With this map, known as the 'Google maps of business in Spain' journalists, investors or other parties get a clear overview of the network of administrators.

infoempresa.com

Spain

Infoempresa.com

"Access information on Spanish companies and their administrators"

infoempresa.com

Third, the most popular business model to add value through the re-use of Open Data is to transform it into knowledge and services. Beyond the collection and curation of data, this aims at offering insights based on the data collected that are meaningful for decision making. This business model matches the archetype 'Developers/Enrichers' from the Data Value Chain, in which value is created by combining Open Data with certain skills in the field of data analytics or development. The role of Open Data for these models varies from providing 'extra contextualisation' to being a core element. In order to transform the services into a product, the interviewed companies sell "Software as a service", offer subscriptions or provide an Open Data based service at the client. As these companies sell services in which they incorporate Open Data, their range of possible business models is much more diverse than those presented above, ranging from the applications of **SmartAppCity** to the drones of **Canard Drones**. With the versatility and plurality of Open Data, it can be used in all sectors.

Ireland

Wripl Technologies

"Cognitive Content Technologies"

wripl.com

The Irish company **Wripl Technologies** is a good example of how by combining skills in computer science with Open Data, new technologies can be developed. This company, which is the result of the PhD conducted by its founder, provides services which can analyse and predict the reception of certain types of online content. With these services, the monitoring of visitors can be analysed and the content production be adapted to fit the need of the visitors. Both of these services are based on the semantic analysis of Linked Open Data on encyclopaedia. With the, digitally accessible, broadness of information that is stored in these databases Wripl can automatically detect the nature of the content and what is related to the subject. Via several algorithms it is then possible to predict what other content could be of interest of the visitors of a site. By providing more content which is fitted to the interest of the readers, readers are stimulated to visit sites frequently and stay longer on the concerned websites. Wripl is steadily expanding their client base by offering potential clients a three month trial period during which they experience the services before entering a contract. Wripl is part of the EU - H2020 Project FREME.

While most of the organisations interviewed transform the Open Data available into actionable intelligence by the use of automated processes, it can also be used as input for stories. This can be seen in the work of the Spanish company Peace Meetings. Another way to use Open Data as a complementary aspect in a business model is by combining it with physical products such as sensors. This way an entire ecosystem can be offered, from the tools to collect data to the analytical services, as can be seen in the work of SmartUp Cities.

The Spanish company **SmartUp Cities** offers Smart City solutions. To help cities become smart, this company not only provides sensors to municipalities by which they can collect data but also assistances in processing and analysing the acquired data into actionable intelligence. In this business model, it is in the interest of both the supplier and the customer to use Open Data as it brings down the price of data collection.

SMARTUP CITIES

Spain

SmartUp Cities

"Smart City IoT Solutions for the Future Cities"

smartupcities.com

Peace Meetings
Spain

Peace Meetings

"Communication for Development"

menplanner.org

Peace Meetings uses Open Data as input for stories to catalyse and foster debates on Peace keeping. With these stories, the company is working on the dissemination of the Culture of Peace to find peaceful solutions for social issues. This is done primarily on behalf of municipalities as extracurricular education.

The fourth and final business models distinguished by the interviewees are those without a specific revenue model. These are primarily applied by organisations which are funded by the public sector or which are non-profit organisations such as Wikimedia. Despite the fact these organisations are not creating direct financial revenue, they are creating value for the society at large with their use of Open Data. For example the Swedish organisation **eGovlab**, uses Open Data to address civil challenges and to accelerate new economic activities. Another example of a non-profit organisation who leverages the benefits of Open Data for value creation is **CRIGE-PACA**.

With four tracks, **eGovlab** brings together various stakeholders and means to jointly develop solutions for civil challenges. When successful, the solutions can then be transformed into an economic opportunity for the private sector. This example shows that even while the concerned organisation is not for profit, their work with Open Data can create value for both the society as for other economic ventures.

egovlab
Sweden

eGovlab

"Applying the future of inclusive governance"

egovlab.eu

CRIGE
France

CRIGE-PACA

"The Geographical Portal of Provence-Alpes-Côte d'Azur"

crige-paca.org

CRIGE-PACA, created by the local authorities of the French region Provence-Alpes-Côte d'Azur, collects, centralises and standardises geo-spatial data collected by the local geographical departments. This facilitates the exchange of data among these departments and reduces the financial and effort costs. Furthermore, CRIGE-PACA has a facilitating role in shared projects but as well with external relations. This can lead to the inclusion of national public sector bodies but also by bringing in private sector organisations.

All four types of business models use Open Data in different ways to create value, either for themselves, their clients or for the society at large. Whilst conducting the interviews, several organisations indicated to still be in the 'bootstrapping' phase, investing all internal cash flow in the development of the business, to accelerate their transformation from start-up to business. Although this may have had an effect on certain elements of their business model, for example applying a freemium model to quickly expand their client-base, the examples provided above underline the diversity in models for value creation.

4.2.4 Web shops and websites are the most important distribution mechanisms

The organisations use various mechanisms to distribute their added value, although the most popular means to do so is via a web shop and/or website, as shown by Figure 45. Respectively 12% and 17% of the organisations (also) make use of other distribution mechanisms. For example, three organisations used a completely automated mean to distribute, by using an API or repository. Two organisations use Open Data for journalistic purposes and therefore distribute their work by media channels.

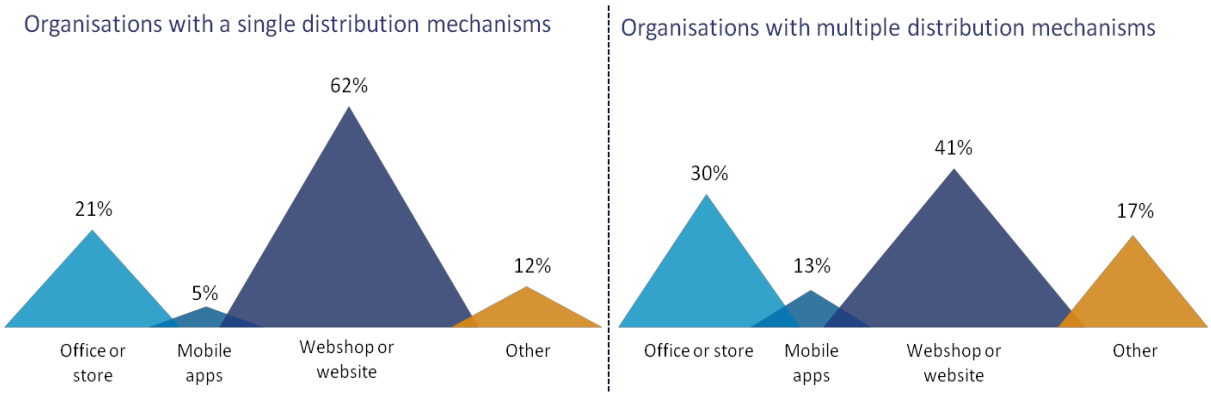


Figure 45: Distribution mechanisms (in percentages)

The Bulgarian company **Mozaika** mentioned a webshop or website as the main distribution mechanism. Mozaika combines data science, natural interfaces and human insights to provide research and solutions to enable productivity and user experience enhancement. Not only do they use Open Data for their analyses, for example by linking it to other datasets, they also intend to publish Open Data as Linked open Data. Within their databases, Mozaika currently offers data infrastructures in multiple verticals, ranging from Business Information to Human Resource Management or for preservation and presentation of cultural heritage, and targets Earth Observation. In all of these services Open Data plays a central element and where possible the collected data is aggregated and shared.

Bulgaria

Mozaika

"The Humanizing Technology Lab"

mozajka.co

4.3 The Open Data capacity of organisations

To derive the potential value from Open Data, organisations need resources. Employees are often the most important resource for organisations. As shown by Figure 46, both concerning the number of employed and external FTE's (e.g. consultants, temporary employees) most organisations indicated to have 0-5 colleagues or employees. The same answer was predominant when asked about the number of employees directly working on Open Data; with 85% of the organisations indicating 0-5 employees. Two organisations indicated to employ more than 1,000 employees. One is a governmental organisation and the other is part of a larger group, which distorts the figures[24]. The respondent with the biggest dedicated group of employees working with Open Data is BBVA Data & Analytics, the data-based analytical company of the BBVA Financial Group.

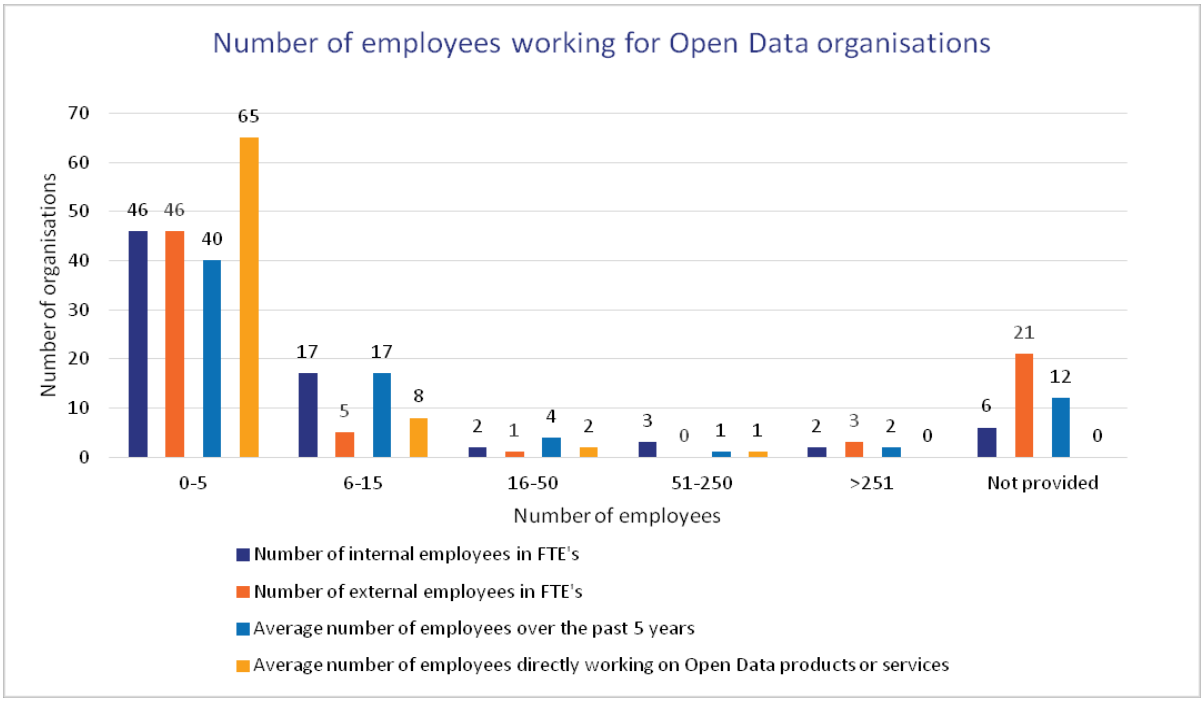


Figure 46: Number of internal and external employees with a full time equivalent workload (in absolute numbers)

Although most of the organisations that took part in the survey are relatively small, most of them foresee to grow in the future. As can be seen in Figure 47, over three-quarters of the organisations expects to hire new staff. A possible explanation is the fact most of the respondents are currently in a start-up phase. Their hiring forecasts match the growth predictions described earlier in this chapter. Once their product has matured, they expect to grow exponentially which will result in more employees.

[24] These are respectively the City of Montréal and the BBVA financial group

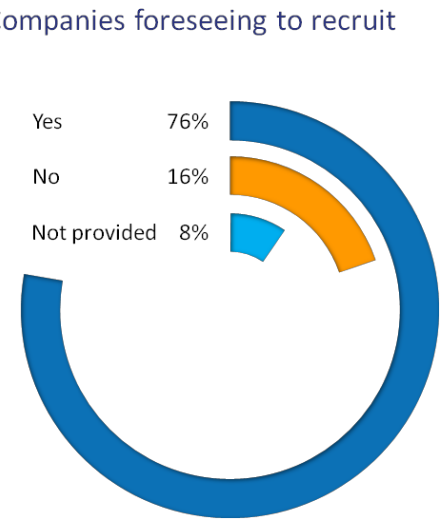


Figure 47: Companies foreseeing to recruit (in percentages)

Whereas the vast majority of the organisations included in the survey are relatively small, this does not mean that larger corporations are ignoring Open Data[25]. It may underline the fact that larger organisations may not wish to openly communicate about how they use Open Data. In addition, it gives an indication that relatively small and young organisations are more eager to reap the benefits of Open Data and to communicate about them. Multiple explanations were mentioned by the interviewees to justify the strong presence of smaller organisations. For example smaller and young organisations (start-ups) were argued to be more flexible than larger ones, enabling them to quicker act upon new opportunities such as Open Data. For these type of organisations, the limited costs of using Open Data was also brought forward as explanation.

Combining insights from the survey, interviews and the IODC temperature check, this chapter described what business models organisations have developed to make the use of Open Data economically viable for them. These insights help get a better understanding how organisations transform the freely available Open Data (economic) value and what kind and form of data they need for this. For most of the organisations their turnover is only marginally related to the use of Open Data. A possible explanation could be that they primarily offer certain services (e.g. software or research) for which Open Data can be used as an element such as research services or Internet of Things (IoT) solutions. This means that although the organisations use Open Data, the turnover is generated primarily by the service and not by the use of the data itself. Bringing back into account the relative youth of the organisations, the relatively low share of Open Data turnover is not surprising. During the start-up phase most organisations are still heavily investing in product development in times of a low turnover. The next chapter will delve into the particular skills that are needed to transform Open Data into economic value.

[25] European Data Portal, 2015, Digital Transformation and Open Data

5. What skills are needed to work with Open Data?

Working with data requires a number of skills, such as technical skills to combine multiple datasets, or business skills to promote and market new services using data. This chapter explores the diversity in profiles and skills companies are seeking to work with Open Data. The chapter concludes on measures taken by different organisations to keep their employee's skills up to date.

5.1 Profiles companies are seeking to hire

Depending on the nature of the organisations' work with Open Data, different skills may be required. These profiles also evolve over time as the organisations grow and diversify their activities. The companies were asked to characterise the profiles they would be interested to hire. As shown by Figure 48, the most sought for profile is data scientist; a profile typically focused on the development of the internal functioning of the service or product.

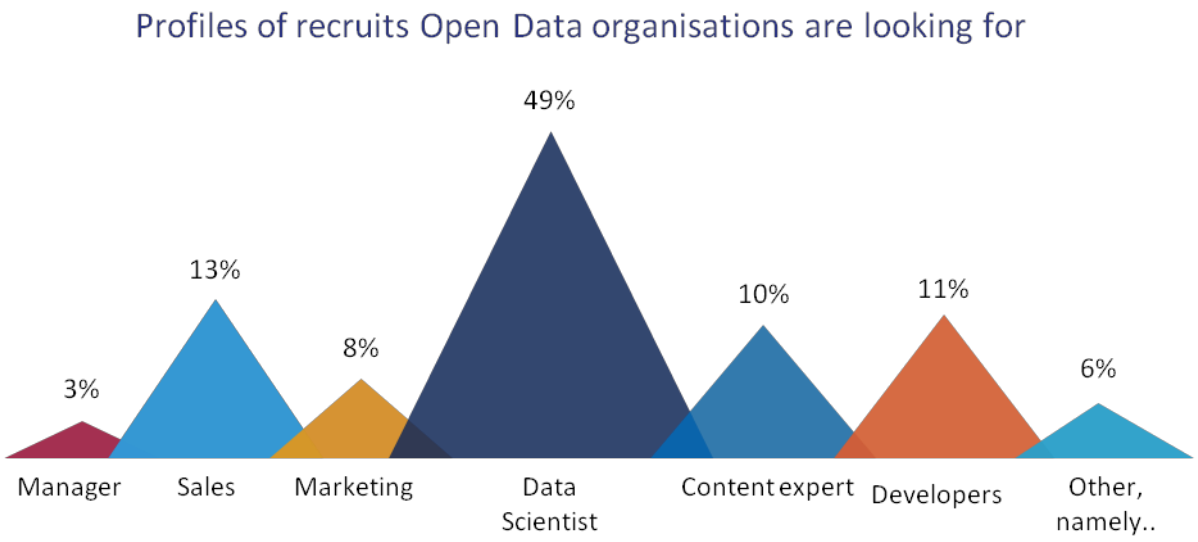


Figure 48: Profiles respondents are looking for (in percentage)

When the founding year of the organisations is taken into account, a strong preference for data scientists can be identified, particularly for the younger organisations (founded in or later than 2011). Of this group of respondents, representing 58% of the entire survey, 61% indicate to look for a data scientists. For the organisations founded before 2011, the profiles sought for are less specific. A possible explanation for this strong preference in the younger cohort is that most of these organisations can be classified as start-ups which are (still) developing their products and services. A similar, though less strong, relation can be identified between the size of the organisation and the profile preference. For the organisations with on average less than 10 employees over the last five years, 50% indicate to look for a data scientist. This again would follow the argumentation of a strong representation of young and small start-ups which are still in the process of product development.

What competences are expected from a data scientist? The European e-Competence Framework (e-CF) provides a standardised way of describing competences that are required for specific ICT jobs. It uses common language regarding competences, skills and proficiency levels that can be understood

across Europe.[26] The role of 'data scientist' was also described using the e-Competence Framework. The competences are shown below.

e-Competence	Level[i]
A.3. Business Plan Development	Level 5
A.8. Application design	Level 2
A.7. Technology Trend Monitoring	Level 2
A.9. Innovating	Level 5
B.1. Application development	Level 3
B.4. Solution deployment	Level 3
B.5. Documentation Production	Level 3
D.10. Information and Knowledge Management	Level 5
D.11. Needs Identification	Level 4
D.12. Digital marketing	Level 3
E.1. Forecast Development	Level 4
E.3. Risk management	Level 2
E.4. Relationship management	Level 3

Figure 49: required e-competences for a data scientist[27]

Interestingly enough, 13% respectively 8% of the organisations underlined they were looking for people with a sales or marketing profile in order to promote their solutions further and generate additional revenue and or visibility. 10% seek further content experts in order to analyse the data further and derive insight for either their direct use in the creation of new services, or do deliver further analysis to their clients. Finally, additional developers and programmers are equally sought for. As a conclusion even if the data scientist profile ranks as the most demanded profile, by definition this type of profile in itself brings together a diversity of skills.

For **Smartbow**, the Austrian company behind the Open Data platform DADAFI, the combination of technical skills and domain knowledge is key to the success of their product. By making the ear tags of dairy cattle intelligent, they enable farmers to monitor their cattle's behaviour and developments such as diseases, heat detection or the real time location of individual cows. By automatically collecting the data from the cows, this helps the farmer save time and money in taking care of the cows. To maximise the potential information found in the data, the company shares the collected and aggregated data over the DADAFI platform for other developers to explore its potential and diversify the farmers' choice in applications to use.

[26] The proficiency levels (e-1 to e-5) of each e-competence range from e-1 being an associate to e-5 being a principal.

[27] http://e-cf.nl/wp-content/uploads/2013/03/Onderzoeksverslag-e-CF-profielen-ECP.pdf

The need for data scientists was further underlined in the interviews during which it was mentioned most frequently. In these conversations, the profile included a variety of skills such as programming, analytical and statistical skills. Although not all respondents were under the impression working with Open Data was different than working with other types of data, half of the interviewees did notice a difference working with Open Data. For this group, working with Open Data brings certain particularities not experienced in working with proprietary data which have to be taken into account by a potential employee. The employee needs to understand for example the concept of Open Data as a shared resource, but also needs the skills to contextualise or refine the data to his needs.

Whilst technical skills are the most popular individual particular skill, most of the respondents indicated to look for more than one particular skill in a potential employee. When delving into the details, respondents considered data scientist profiles as bringing a mix of skills such as domain knowledge, statistical skills, as well as analytical skills, business insight and communication skills. Despite the diversity in skills sought for, businesses underlined the technical skills they expected such profiles to master. Figure 50 below provides a distribution of these skills.

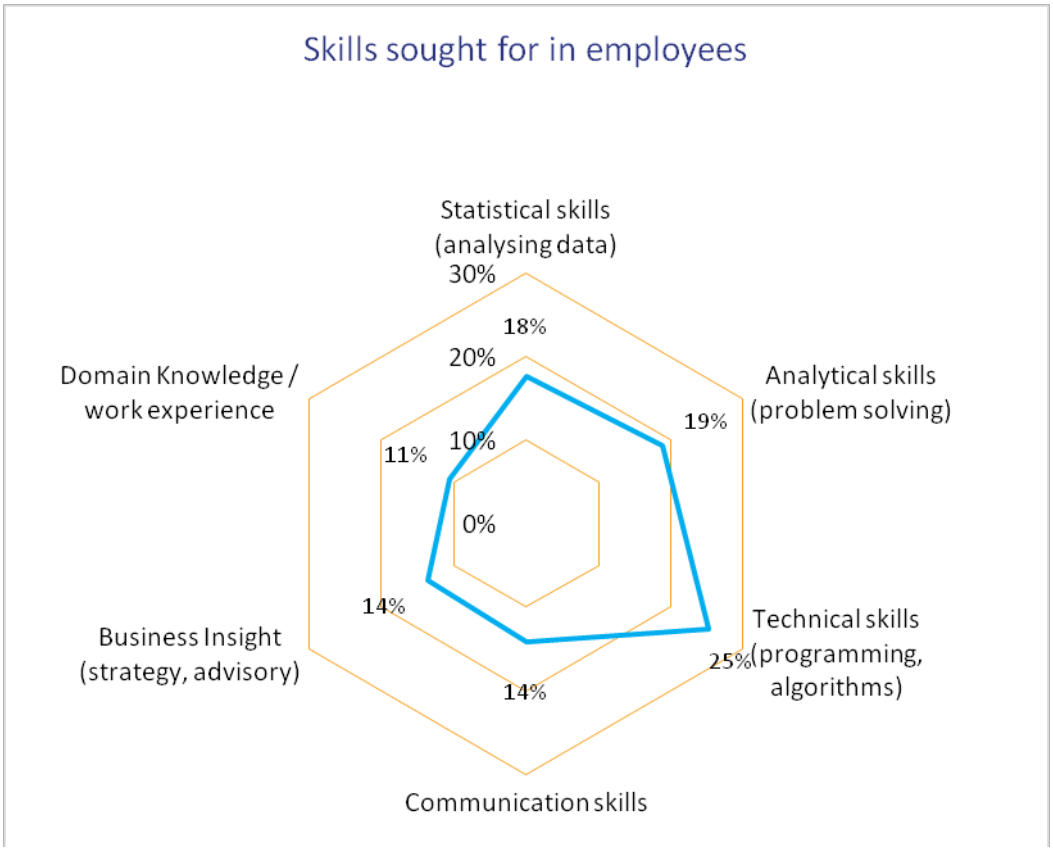


Figure 50: Skills sought for in employees (in percentages)

When the combinations of sought after skills were examined closer by the use of a correlation matrix, a number of interesting findings were identified as can be seen in Figure 52. The combinations of skills either tend to focus on skills needed for the development of a product (back-office skills also known as 'hard skills') or at the development of a business with the product (front office or 'soft' skills).

This plays in tune with an earlier report published by the European Data Portal on the type of skills needed to work with Open Data in which a typology was developed between hard and soft skills. [28]

[28] European Data Portal, 2015, E-skills and Open Data

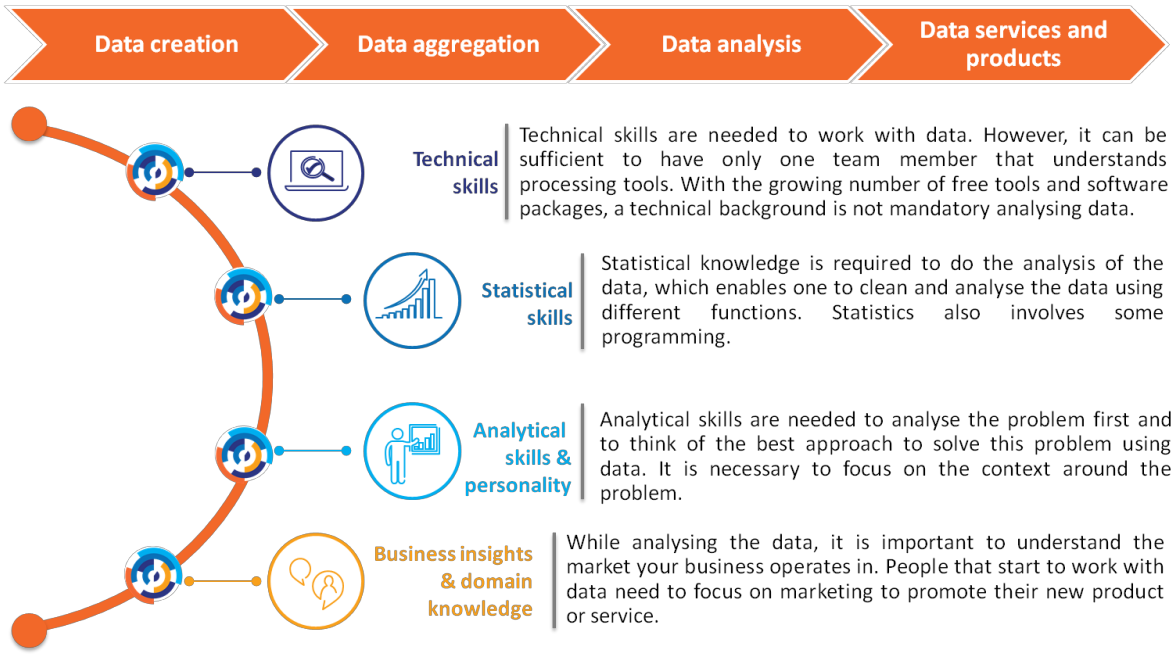


Figure 51: Skills required to work with Open Data

Whereas technical and statistical skills can be considered as hard skills, analytical skills, personality, problem solving, storytelling, collaboration, curiosity, communication and creativity are considered soft skills. When building a correlation matrix clustering the skills, there is a clear preference for hard skills. As can be seen in Figure 52, the strongest correlation is found between the statistical, analytical and technical skills. The least mentioned combination is communication combined with technical skills, although communication skills themselves was frequently mentioned as a necessary when seeking to enter a market.

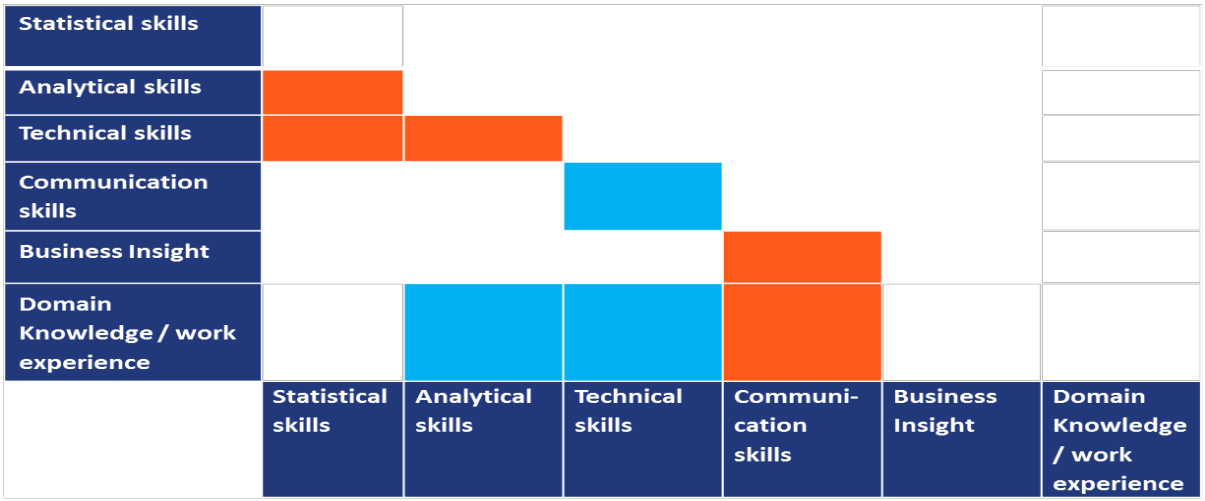


Figure 52: Correlation matrix of skills (correlations > 0.1 are marked in orange, correlations < -0.1 in blue)

To further confirm this, the respondents were also asked to rank these skills. The dispersion of skills ranked with the highest value are displayed in Figure 53, which reconfirm the importance of both technical and statistical skills but place analytical skills as the highest valued skill.

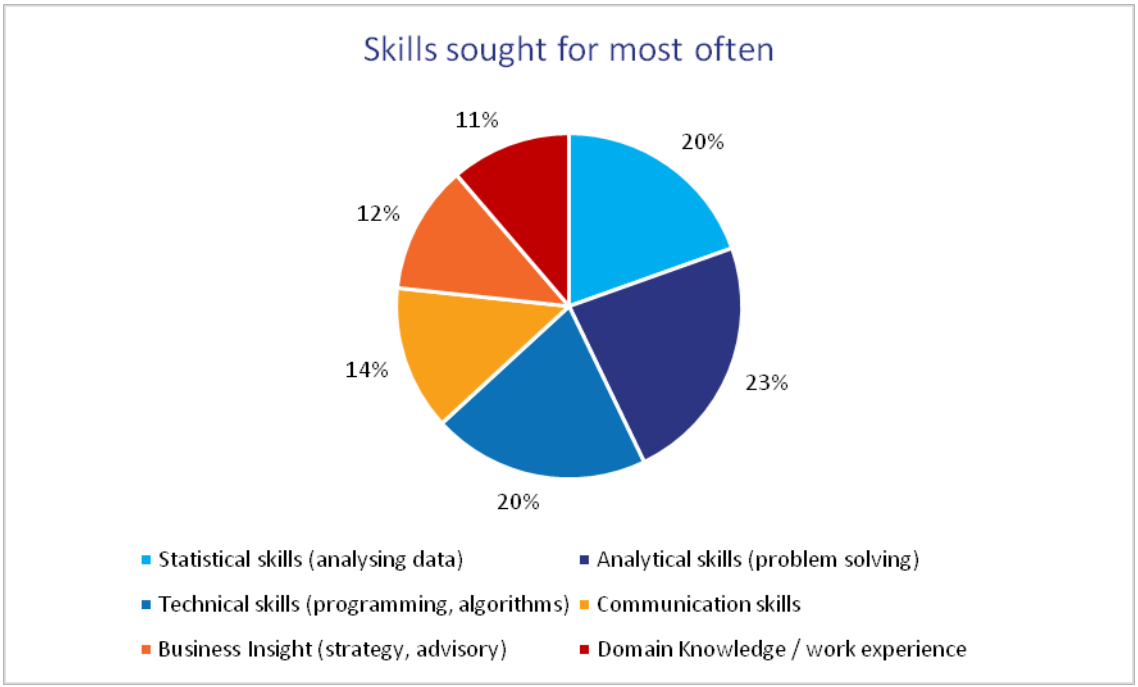


Figure 53: Skills sought for most often (in percentages)

5.2 How to acquire these skills?

76% of the interviewed organisations indicated to be active in the acquisition or up-keeping of the skills of the people working in their organisations. Based on the variety and dynamic landscape of Open Data, respondents frequently mentioned however that by focusing on traditional curriculum is not necessarily the most effective way to find suitable candidates. To overcome this, many of the organisations use self-training and training on the job for the employees to acquire further skills. 30% of these companies indicated to provide limited or tailored education and training to their employees, motivated either by financial reasons related to the maturity of their company or by providing only very domain-specific training. For the organisations that work with larger communities of developers and volunteers or the non-typical organisation structures, videos and eLearning are to be an effective way to disseminate the necessary skills. For example, **AppChallenge** indicated to provide training to its community using its YouTube channel.

Although the respondents are predominantly positive about finding candidates with the required skills in the future, the opinions and visions are mixed. Overall, the 'positive group' sees a growing demand, maybe even a hype, for data scientists which can create a short term shortage. Adjacent to the 'hype' they do expect to see more people obtaining degrees in this field which will offer a larger supply of data scientists. However, business knowledge and the capacity to contextualise data will be necessary as well, underlining the need to build mixed teams and collaborate within and across organisations.

If the pool of available data scientists does not match the demand in the future, a potential barrier arises for the further uptake of the re-use of Open Data. This and other barriers are presented in the next chapter, which touches upon barriers, success factors and recommendations.

6. The way forward

The previous chapters provided insights into the economic value of Open Data and how organisations benefit from this value. However, there are still some barriers that prevent organisations to fully reap all of the potential benefits of Open Data. These barriers have a negative impact on the continued re-use of Open Data. This chapter dives deeper into the specific issues by addressing the different kinds of barriers encountered by organisations working with Open Data and how to face them. In addition to the challenges and the barriers identified, a number of success factors on how to work with data are described. To add to the prospects offered by Open Data, together with the organisations that have participated in the survey and interviews, a wish list has been established.

6.1 The main barriers encountered by businesses

Barriers can be diverse, ranging from technical to business concerns or dependencies on external actors. For many of these barriers, the respondents indicated that they are intrinsic for working with Open Data and accept them as 'part of the game', perhaps even as trade-offs for obtaining data without any (direct) costs. The most recurrent barriers identified for working with Open Data are:

- A poor quality Open Data,
- A lack of standardisation or heterogeneity,
- Difficulties in obtaining the data with the right information (metadata) for the purpose or its usability.

The **low quality of Open Data** was perceived both in the data itself as well as in the accompanying metadata. In both cases, respondents recurrently mentioned examples of Open Data they found in which either the indicators used within the set or the metadata describing the set did not correctly represent the data. Without correct descriptions of the data, the user of the Open Data has to make certain interpretations leading to more work and a greater risk of misinterpretation of the data. Multiple explanations were brought forward by the interviewees as reasons for the low quality of Open Data. The most transcending explanation was that they perceive the publication of Open Data as a low priority for the data provider, creating limited stimulus for these bodies to publish high quality data. Although most interviewees experienced a positive attitude of the providers when the Open Data policies were developed, in practice they do not always see the (political) promises being translated into practice.

Spain

ISEA

"Technological development, innovation and entrepreneurship"

iseamcc.net

Having a good level of standardisation will not only improve the discoverability of Open Data, it will also enhance how efficiently the data can be exploited. This is in particular the case when multiple data sets are combined with one another as the Spanish company **ISEA** is doing with their project TENTU. With this project, ISEA offers users the opportunity to have news and information aggregated for them in a personalised and thematic manner. As this requires the processing of textual data, it is important the descriptors are kept the same or coherent, event when multiple languages are combined.

While most respondents have the capacity to clean and refine low quality data themselves, the two barriers of experiencing a **stark heterogeneity and low usability** in Open Data persists. These barriers create a vicious circle for the re-users as the one barrier prohibits the development of a solution for the other. In particular heterogeneity - a lack of standardisation of the data, variety in platforms to access data or low consistency over time - restricts the opportunity for users to develop permanent solutions to overcome these types of barriers. Another frequently mentioned barrier is a **poor discoverability** of the exact type of Open Data users are looking for, a barrier which is related to the aforementioned low levels of quality in the descriptions of the sets and plurality of platforms. When these descriptions are not specific enough or the set is disseminated on a particular platform, it is hard for the user to get the data they need even when it is published.

NETAGE

Netherlands

Netage

"Smart data for smarter fire protection"

netage.nl

A good example of how having more data available could improve not only the functioning of a company using Open Data, but also serve a greater societal function can be seen in the work of **Netage**. This company uses Open Data to help fire departments improve their operational processes, both in the preparation phase and in emergency situations. For this work, having more Open Data available could lead to better analyses to identify the best location of a fire station. Having more Open Data available on a granular level, for example the economic activity within certain buildings, enables fire-fighters to make a better judgement of a situation.

Barriers
for the re-use of Open Data

Quality of Open Data

How to ensure and apply quality standards?

Lack of Standardisation

How to develop interoperable formats, vocabularies and structures?

Availability of Open Data

How to make Open Data more easily accessible?

Incorrect metadata

How to ensure high-quality and standardised metadata?

In order to cope with data heterogeneity, the **Hammer Project** developed a new approach to deal with the variety of portals it collects data from. With growing awareness of Open Data and an increasing number of publishers, the amount of Open Data available is growing by the day. To help clients extract Open Data they need from the pool of information, the Italian company Hammer Project has developed an algorithm, which automatically selects and retrieves the relevant data. Once retrieved, the Data is translated into a suitable format for analysis, statistics or the development of a new application. With their services, the Hammer Project wants to contribute to the use of (big) Open Data for research projects by helping clients become more efficient in the first steps of the Open Data Value Chain, the steps where the information is selected and retrieved.

Italy

Hammer Project

"search and visualize OPEN DATA"

hammer-project.com

These barriers are largely supported by the outcome of the Temperature Check. Both public sector and NGO representatives mentioned the quality of data sets to be the main challenge when working with Open Data. Especially public sector representatives perceive quality of data to be the main challenge as this was mentioned by 68% of these type of respondents, followed by respectively the availability of data (39%), and standardisation (32%). 51% of the NGO representatives indicated the quality of data sets to be the main challenge when working with Open Data, followed by standardisation (45%) and availability (42%). Over half of the private sector representatives (61%) indicated the availability of data sets to be the main challenge, followed by quality concerns (51%) and standardisation (42%). A clear overview of these barriers is shown in Figure 54.

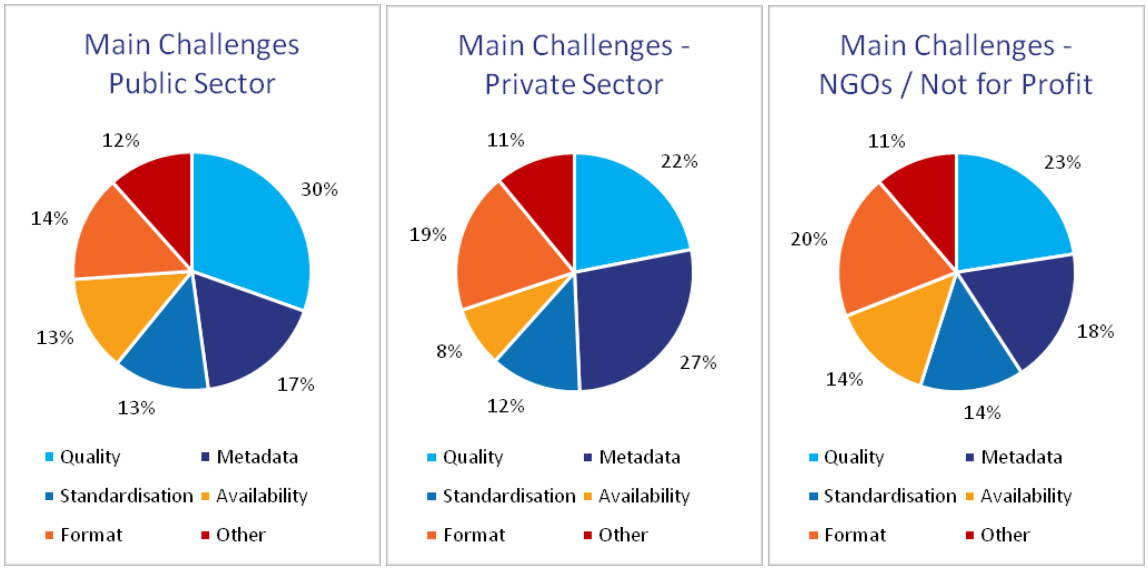


Figure 54: Main Challenges - Public Sector - Private Sector - NGOs/Not for profit IODC Temperature Check (in percentages)

Many organisations consider these barriers to be intrinsic to working with Open Data. For them it is the 'learning by doing' by which they have improved their work and enabled them to deal with these type of barriers more efficient. As they cannot fix the barriers which are created by the data providers, they do indicate to spend more time on quality assessments before further using the found data. While most of the interviewees were confident these barriers do not or will not create irreconcilable issues, one interviewee indicated not to have had started his Open Data business if he had been aware of the barriers before founding his business. From a business perspective, the interviewees mentioned that in hindsight the client's perspective should have played a bigger role when founding the organisation. For some, a lack of client orientation at the start led to the development of products which did not exactly fit the need of the actual client and could have been prevented. At the same time, some interviewees underline that this also led to unforeseen innovations.

A good example of how the scarcity of Open Data can be related to the domain of activity is demonstrated by **Doctrine.fr**. This French company uses Open Data on French and European case law to provide legal analytics. Their database of case law builds on this is accessible to all, thanks to their paid services lawyers are offered highly efficient tools to access relevant case law. Although Doctrine.fr prefers to work with Open Data, only a marginal number of case law is currently published as Open Data which hinders further growth of the company.

Doctrine.fr

France

Doctrine.fr

"Le big data au service du droit"

doctrine.fr

Open Data is often perceived as being a technical topic. As shown when assessing the skills used and sought for by companies re-using Open Data, hard skills appear to be dominant. In this respect, working with Open Data can indeed require a number of technical skills at different stages of the data value chain. Overall, the interviewees were under the impression that even while technical aspects could arise, they are all surmountable. With their own (often technical) background and the continuous progress witnessed in the tools being developed to work with data, 30% of the interviewees said there are no technical aspects which could create major problems for their work. Other interviewees mentioned certain technical aspects to be a barrier, such as the need to be familiar with APIs in order to acquire the data. Additional barriers mentioned concern for instance the need for need for a sufficient hardware infrastructure in order to deal with the big volumes of data, a need some fulfilled by using external processing capacity with cloud services. In order to build such an infrastructure, the organisation needs adequate (human) capacities in programming and database management.

The need of technical knowhow on APIs or sufficient hardware to deal with big volumes of data was for instance mentioned by the Dutch company **Kenedict**. This company makes use of Open Data to provide additional context to the innovation networks of its clients. By visualising and analysing data on market environments, Kenedict Innovation Analytics provides tangible and meaningful insights for organisations which focused on innovation. These insights can lead to new collaborations, adaptations in their strategies or help them adopt an Open Innovation practice. By using Open Data, Kenedict provides a more contextualised advice to its clients, for example by interactively visualizing networks of collaborating organizations based on EU research funding data. When this company started to include Open Data in their analyses, it had to familiarise itself with the role of APIs as Open Data channels. The company not only applies their knowledge of Open Data visualisations in the projects for its clients, they also enable others to experience the benefit of visualisations with the application Kenelyze. On this platform, the company facilitates people to upload their own data to visualise and analyse the networks residing within any dataset.

6.2 Success factors for working with Open Data

If data re-users have developed workarounds to address the bottlenecks they face, what are the underlying key success factors that they have identified so far? The key success factors can be grouped into two key themes corresponding to external and internal factors, which can be identified based on whether the factor is depending on the data itself or the processing with the data.

External factors: **getting the right Open Data.** Here, re-users are dependent on what and how data providers make and deliver their Open Data to the re-users. For some of the re-users, this dependency is of such importance that it is critical to the economic viability and sustainability of their organisation. A dependency like this can be avoided by diversifying the required data or by looking at other channels to acquire it. Having a particular type or form of data as a success factor is also contingent to the specialisation of the organisation. Whereas an organisation which is predominantly active in one domain or sector has a higher chance of being dependent on the provision of that specific data, an organisation working on the processing of data regardless of its kind can easier diversify their data need.

The survey also touched upon the type of Open Data or language the organisations need or prefer. Within the survey, organisations were also asked to determine which data is crucial for the success of the business activities of the organisations, they were asked to identify the critical sources of data for

their respective organisation. As shown in Figure 55, the two most crucial elements for the organisations are receiving data of high quality and being able to rely on the fact that the used data sets are published systematically and continuously.

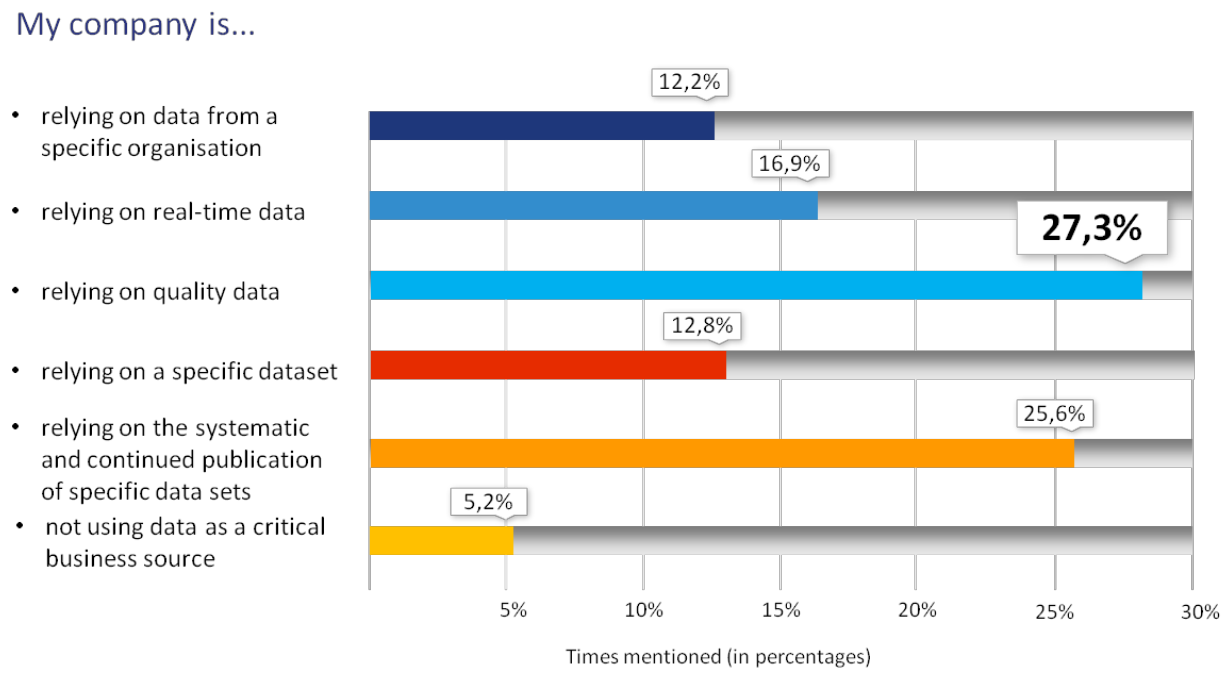


Figure 55: Critical elements for organisations (in percentages)

For 41% of the respondents, satisfaction of these criteria (quality and consistency) were considered to be critical for a successful operation of their business, indicating a strong correlation between the quality and consistency. The importance of consistency for Open Data, both in content and frequency, was illustrated by one interviewee who even indicated to stop using Open Data if this would not improve. Although the other forms of data are less frequently mentioned, it is clear that the organisations have a certain amount of requirements for the provision of Open Data for them to be able to work with.

Internal factors: working the Open Data. The internal factors relate to how the organisations transform the data into a product. For this to be successful, the organisations has to **transform the data into knowledge** and use this to solve a business challenge. As the data itself is very generic, the organisation has to transform this into knowledge or actionable intelligence. This is of particular importance as the used (Open) Data itself is not the asset, as it is open and freely available to everyone. The asset of the Open Data organisation and thereby its added value is how it is able to transform the data into a workable knowledge. This knowledge can be used within the organisation itself or to deliver insights to external clients.

SmartAppCity has quite a different perspective on how they use Open Data in their company as instead of focussing on one particular subject they aggregate and collect all Open Data. By bringing all the available data and information together into one application, they can create a comprehensive application which is interesting for both visitors and citizens of a city to use. Due to the broad array of subjects covered in this application, it becomes interesting for both public and private bodies to provide it with information. To further encourage and highlight the collaboration between local governments and businesses, the application aims to function as a public-private partnership. In this partnership, the development costs of the application are borne by the municipality and the maintenance costs by the businesses for which it is an additional communication channel.

SMARTAPPCITY
Spain

SmartAppCity

"All the city in your hand"

smartappcity.com

The second internal factor to be successful as Open Data organisation is to **solve a (client's) business challenge**. In line with above, this is about using data to solve a given problem or to generate demand for a given service or product. For some interviewees, their organisations initially focused more on the product itself than on what problem the client would like to see solved. To successfully market the product however, the client – or the organisation itself – should not fit to the product but the product should be aligned with business need, offering a 'unique' value to the clients

Language was also seen as a factor which could influence the development of a successful Open Data organisations. English seems to be the *lingua franca* for most of our respondents as over 41% indicated to prefer data in this language (*Figure 56*). On average, the respondents indicated to use data in more than one language. The low average of preferred languages could be partly explained by the number of respondents who offer their services locally in a specific country or language area.

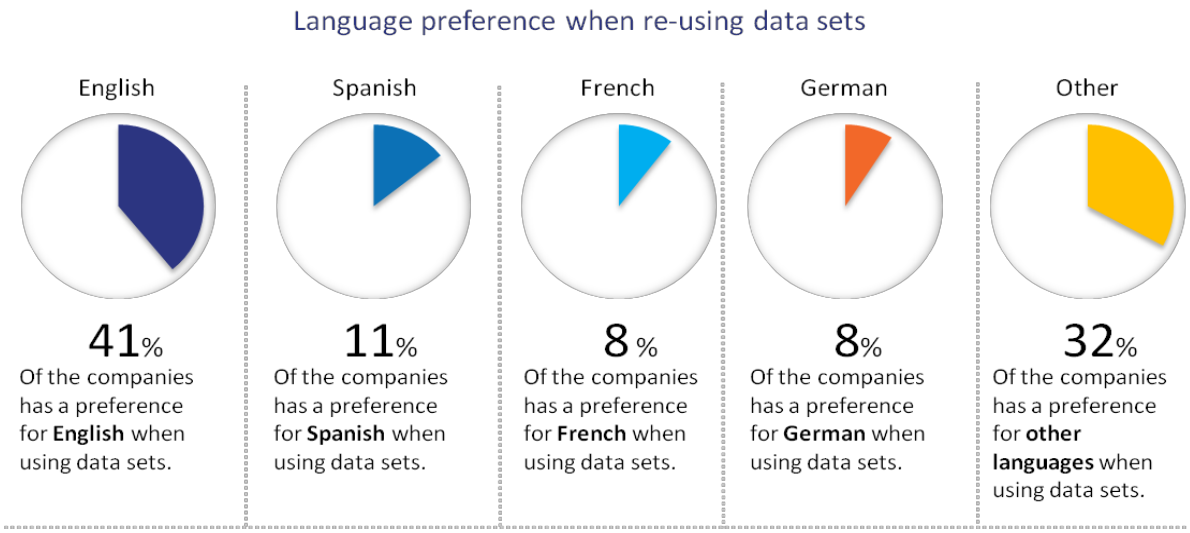


Figure 56: Language preference when re-using data sets (in percentages)

However, when asked about whether language did matter, the predominance of English was less apparent. The majority of the respondents answered affirmative (Figure 57).[29]

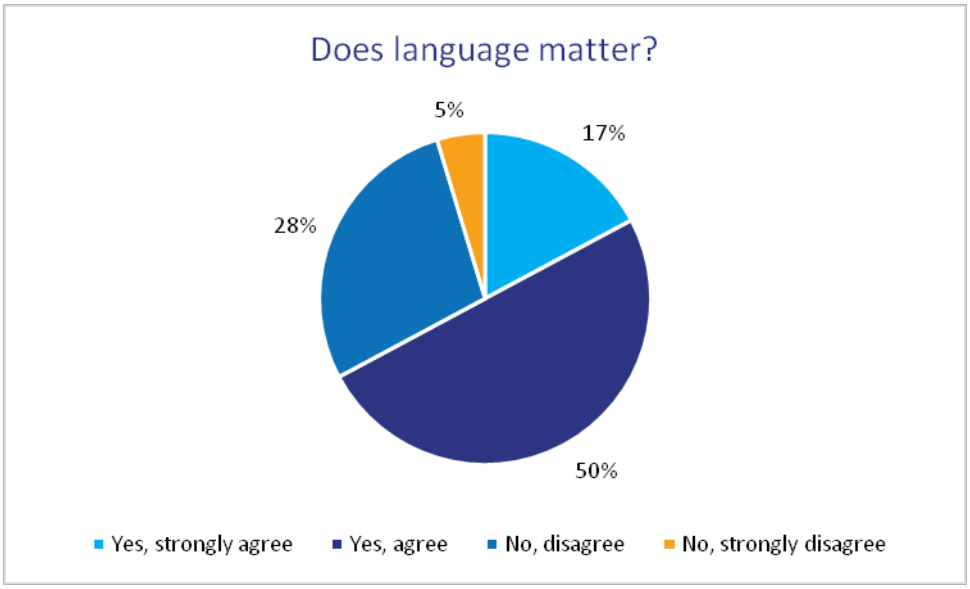


Figure 57: Does language matter? (in percentages)

While several interviewees mentioned domain specific motivations as cultural or judicial reasons to motivate the use of data in a specific language, others mentioned that having to translate the data would lead to a loss of quality. Some respondents also mentioned to use machine translation to overcome this and homogenise it into one language. This explains why certain data aggregators prefer one language while using data from multiple language as automated translation enables them to collect data in multiple languages but to process it in one particular language.

[29] In total there were 43 respondents who indicated to agree or strongly agree with this statement

SUCCESS factors
for working with Open Data

External factors - Getting the right Open Data

Data for job

- What do you need?
- Where to find it?

Quality of the data

- What format?
- What level of detail?

Consistency of the data

- Availability over time
- Consistency and comparability over time

Internal factors - Analysing the data

Transform the data into knowledge

- Identify the information in the data

Add unique value to the data

- Combine with existing data to drive insight

Solve a business challenge for the client

- Work from a client's perspective

6.3 What do organisations using Open Data want for the future?

What do organisations want from Data publishers? A number of barriers stem directly from the publishers themselves, but what other aspects could be improved? 33 organisations provided their unfiltered comments on what they would add to their wish list. Top of the list, the majority of organisations indicated wishing to have access to organisation/trade registries, economic data, legislative or case law data, geospatial data and data on the public administrations. However, for some respondents, the emphasis was not so much on the quantity but rather on the quality of the data. While the current legislation is primarily concerned with which data ought to be released, the interviewees would like to see a stronger description of the form and structure for the release of Open Data.

Slovenia

eVineyard

"The vineyard management software ahead of time"

evineyardapp.com

Although **eVineyard** has particular data sets it would like to see opened up, it is a good example of seeing particularly more homogenised data being made available in uniformly created lists. This company from Slovenia provides a software suite and supporting IoT sensor nodes to help wine producers with the management of their vineyard. To provide the most accurate advice possible to the producers, eVineyard uses a variety of data types such as satellite, weather or agricultural data. While there is a lot of Open Data in these fields already, what would help eVineyard most in expanding rapidly in multiple countries is getting this data at a unified and European level.

Also **Glimworm**, the Dutch company behind the IoT Living Lab in Amsterdam, is under the impression that legislation such as the PSI Directive is a great tool to get more data being shared with the society. However, according to this company it is not only the volume of Open Data which should be improved but also the quality and value of the Open Data. While for certain subjects, this would require particular institutions such as the Chamber of Commerce to open up, the company is also actively contributing to this itself. For example by collecting and sharing high quality Open Data obtained by their IoT nodes throughout the city but also by helping the municipality share and maximise the use their parking data through an Open Data Platform.

Netherlands

Glimworm

"Tomorrow's web today"

glimworm.com

Raising the awareness around Open Data at both data publishers and data re-users is equally a recurrent topic. The present low level of awareness does not encourage the release of good quality Open Data or the spread of use in the private sector. Greater levels of awareness on the Open Data benefits by the society at large, including the private and civic sector, is expected to lead to a higher use of Open Data. When more people use or know about the usage of Open Data, this will lead to a higher political pressure for the public sector to provide (high quality) Open Data.

To explore if the organisations include Open Data in their commercial activities or branding, they were asked if they communicate on this subject with their clients. Of the organisations, 65% indicated to communicate about Open Data with their customers. These organisations have a strong belief in Open Data and use this communication to advocate the release of more Open Data. According to the organisations that do communicate about Open Data, they perceive their clients to have an either poor or mixed understanding of the concept. For many of the respondents, the function and position of the client's main contact person has a big impact on the understanding of Open Data. This ranged from a fear of sharing 'their' data to thinking the whole concept of sharing information without compensation was something for 'hippies'. Both of these findings indicate that there is still a lot of unfamiliarity with the concept of Open Data.

The interviewees are strongly advocating for the further development and dissemination of success stories, in order to raise awareness on Open Data for the society at large. With these stories, Open Data and the use of it becomes more tangible and the value more insightful. In addition to the use cases available on the European Data Portal, there are already a number of initiatives across Europe to collect and share stories of Open Data organisations. In Germany for example the website www.datenwirken.de provides an overview of applications based on Open Data. Also many national Open Data portals provide attention to Open Data applications.

While the improvement on awareness requires a culture change, the other frequently mentioned improvements have a technical character. Two types of technical improvements are distinguished, being centralisation of the data and more standardisation in general. By centralising the access to Open Data, for example as done by the European Data Portal, the discoverability of the various data sets will be improved as users know where to look. Centralisation would further reduce the heterogeneity of platforms, enabling users to build more automated ways to download all the data they need. Many of the interviewees mention centralisation of Open Data as an improvement, which could take place either on a national, European or global level.

Centralisation of access to data will also force the question on standardisation and harmonisation to improve the overall accessibility and aggregation of Open Data. With more standardised data, it is easier for users to mix and combine different datasets into new or specialised databases. Furthermore, standardisation of the (meta-) data would improve discoverability of data as search queries can be re-used in multiple settings. In the line of standardisation, some interviewees suggested a central portal could also play a more active role in terms of harmonisation. The work on the DCAT-Application Profile conducted at the European level is a stepping stone in achieving harmonised metadata profiles across Europe.

7. Conclusion

After having presented the economic value of Open Data at the macro-level, this study set out to explore how this (economic) value created at the microeconomic level. The report presented the variety and versatility in re-use of data and different business models that have been developed at the different steps of the Data Value Chain. Different models exist for companies to derive economic value out of Open Data. In particular businesses that have been running for multiple years use Open Data to enhance **their existing services** or make their own work **more efficient**. Other organisations commercialise Open Data within the **actual product or service** they provide to their client. For most of the organisations, their principal source of revenue is by **the provision of services** to their clients.

Organisations deliver these services in all the steps of the Open Data Value Chain, from facilitating (access to) Open Data to completely transforming Open Data into new services. Being digital natives, these services are predominantly distributed using digital channels such as a web shops or applications. Related to the generally short period in which the organisations are in business, some have started as recent as in 2016, most of the respondents having an annual turnover under € 50,000 with less than five employees. Organisations' turnover can sometimes only marginally relate to the use of Open Data, however for a majority of respondents this represents up to **60% of their turnover**. When looking at the employees working for these organisations, many of them have a background in data science or strong analytical and statistical skills. However, it is apparent that in order to be successful in transforming Open Data, these skills alone are not sufficient. To make the use of Open Data into a successful and marketable product, it is key to have knowledge and insights from a business perspective as well.

Furthermore, the study has provided data providers with more insights on how Open Data can be used, after which they may adapt their data provision strategy accordingly. Room for improvement range from increasing the **volume of data available** as well as focusing on **quality of the data** and **metadata** provided. Clear descriptions are needed in order for re-users to be confident that they can rely on this data. Another aspect relates to the heterogeneity of the data, such as in **format** or **(in-) consistency of the metadata**, and the **dispersion of data among several platforms, interfaces and languages**. These factors hamper automated processes, and provide food for thought for data providers when rethinking their provision strategy.

Recommendations for both the data providers and data users are primarily around **raising awareness**. With a greater knowledge and awareness around Open Data, in particular the value found within it, data providers become more aware what can be done with Open Data. By knowing how 'their' data can be used they could align their publication strategies with this need, resulting in the provision of better and more suitable Open Data for those organisations that can transform it into value. Data re-users therefore also have a part to play by sharing their stories and showing how they have been using Open Data.

Annex I: Methods used

In order to better understand the re-use of Open Data, a methodology consisting of three pillars was developed.

1. Online survey

The survey questionnaire consisted of 57 questions. Most questions were multiple choice, however open sections offered each respondent the opportunity to provide free text to further explain theire responses. Of the 57 questions, 11 questions were mandatory.

In order to avoid misunderstandings on the definition of Open Data, the survey started by providing the following definition:

> *"Open (Government) Data refers to the information collected, produced or paid for by the public bodies (also referred to as Public Sector Information) and made freely available for re-use for any purpose."*

The survey was divided into seven themes and where necessary, the questions were accompanied by an explanation or terminology. The survey was open to the public from 18/02/2016 and ran until 05/06/2018. The full questionnaire can be found in Annex IV. The first grouping of 16 questions addressed company details. This was followed by 7 questions regarding the financial performance of the organisation, both historically and forecasted. A section on the human resources involved in these organisations was also included. After having addressed the organisational structure of the respondents, the following sections addressed the use of Open Data.

Of all the valid entries, a number of details were collected explaining the nature of the organisation, for example who owns the organisation. As can be seen in Figure 58, most of these organisations are either owned by the CEO (29%) or by the management team (26%).

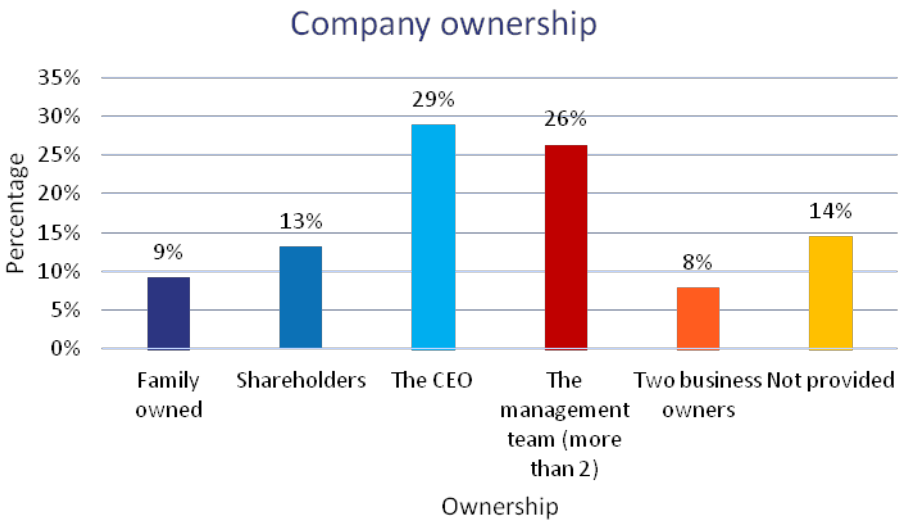


Figure 58: Company ownership (in percentages)

Of these companies, the majority was founded in the early years of the second decade with by far the most in 2014. As can be seen in Figure 59, the outliers are the two organisations which were

founded before 1950. These are the City of Montreal (Canada) and the global financial group BBVA. As the oldest entity to report in the survey, the Open Data re-use of BBVA will be further exemplified in the following chapter. When looking at the nature of the company, either for-profit or non-profit, 60 out of the 76 companies indicated to be a for-profit organisation, a clear indication that value adding organisations are exploring Open Data.

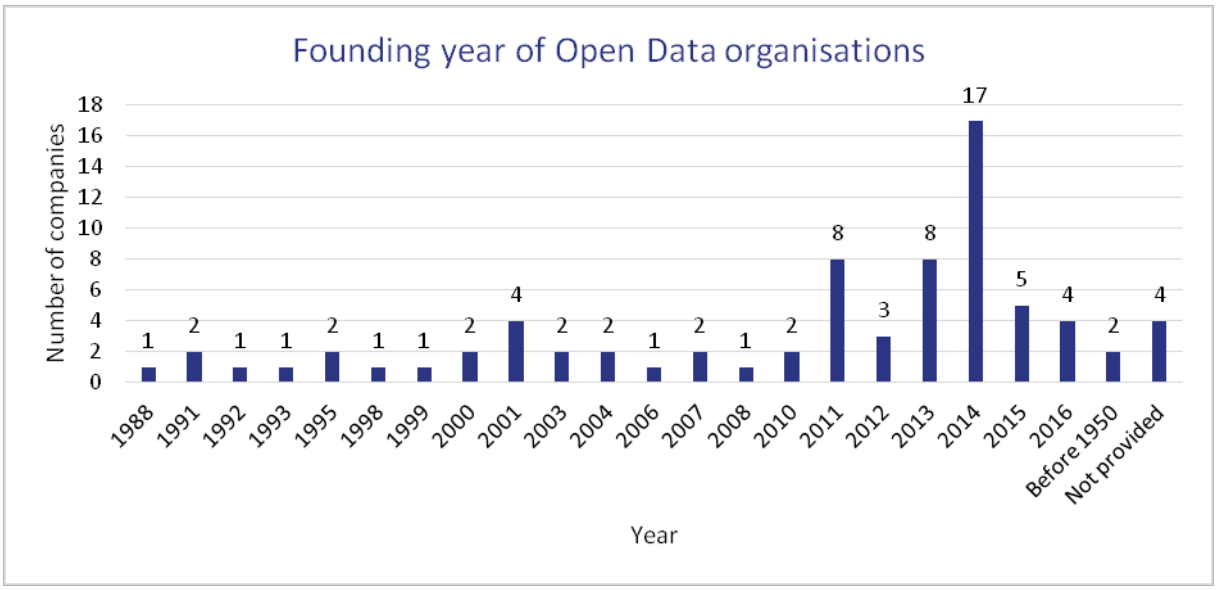


Figure 59: Founding year of all valid entries (in absolute numbers)

2. Interviews

The survey resulted in the receipt of 76 valid respondents. After having excluded the respondents from a non-European country, this number was reduced to 70 respondents. Based on a number of steps, 44 respondents had been invited for an interview of which with 26 an interview had been conducted.

Selection process

The first step of the selection process election was based on the geographical location of the respondents to ensure the broadest spread of respondents possible. As was mentioned previously, not all EU countries and affiliated countries were (equally) represented by the respondents to our survey. This was also reflected in the regional spread of the respondents throughout Europe, with a strong underrepresentation of the Northern and Eastern regions.[30] In total these regions were presented by a mere 7 respondents, while the other two had a cumulative of 63 respondents.

[30] Based on the regional classification as presented by EuroVoc: http://eurovoc.europa.eu/drupal/?q=request&uri=http://eurovoc.europa.eu/100277

Regio	**Country**	**Respondents**
North	*Sweden*	*2*
	Denmark	*1*
East	Bulgaria	2
	Poland	1
	Slovenia	1
South	*Greece*	*3*
	Italy	*8*
	Portugal	*2*
	Spain	*16*
West	Austria	2
	Belgium	1
	France	4
	Germany	8
	Ireland	2
	The Netherlands	4
	United Kingdom	13
Total		**70**

Figure 60: Geographical and regional spread of the EU respondents (in absolute numbers)

For the remaining 63 respondents, a more extensive selection process was carried out to strike a balance in the representation of all countries, region and NACE Sector of the interviewed organisations. This overview can be seen in Figure 61, where the selected respondents from North and East Europe have been marked in green. In this table the total respondents representing a NACE Code or country with equal or less than 2 respondents have been marked in yellow. Taking into account respondents who indicated not to be contacted, this added a total of eight more respondents to the list of respondents to be approached for an interview. In total, fifteen respondents were selected at this point to ensure the broadest selection of countries, regions and NACE code.

The final ten organisations were selected based on a combination of these criteria, ensuring the most balanced spread possible while taking into account the absolute representation of certain countries such as Spain or the United Kingdom. This resulted in two respondents from France, three from Italy, one from the Netherlands, two from Spain and two from the United Kingdom.

NACE	North		East			South			West								
	SE	DK	BG	PO	SL	EL	IT	PO	ES	AU	BE	FR	DE	IE	NL	UK	**Total**
Administrative & support service								1									**1**
Agriculture, forestry & fishing						1											**1**
Arts, entertainment & recreation													1				**1**
Education	1	1				1	1		2		1					1	**8**
Financial and insurance activities									1			1			1		**3**
Human health and social work activities																1	**1**
Information & communication			1		1	1	2		6	1		1	3	2	2	7	**27**
Other service activities							1		2			1	1				**5**
Professional, scientific and technical activities	1		1	1			4		4			1	2		1	4	**19**
Public administration & defence													1				**1**
Real estate activities									1								**1**
Transportation and storage								1		1							**2**
Total	**2**	**1**	**2**	**1**	**1**	**3**	**8**	**2**	**16**	**2**	**1**	**4**	**8**	**2**	**4**	**13**	

Figure 61: Overview of the NACE Codes per country (in absolute numbers)

Methodology interview

In total, 33 respondents were willing to be interviewed to get a better understanding of their Open Data re-use which are presented in Figure 62: Interviewed organisations. One of these respondents preferred not to be referenced in this report resulting in this particular organisation only being included where results are shown anonymously.

Appchallenge	www.appchallenge.net/	United Kingdom
BBVA	www.bbvadata.com/	Spain
Canard Drones	www.canarddrones.com/	Spain
Connemara Programme	www.myconnemara.com	Ireland
Crige PACA	www.crige-paca.org/	France
CropDiagnosis	www.cropdiagnosis.com	Greece
Doctrine.fr	www.doctrine.fr	France
Egovlab	www.egovlab.eu	Sweden
eVineyard	www.evineyardapp.com	Slovenia
Fluicity	www.flui.city	France
Geomanagement	www.levadia.gr	Greece
Glimworm Information Technologies	www.glimworm.com	The Netherlands
GraphDefined GmbH	www.graphdefined.de	Germany
Hammer Project	www.hammer-project.com	Italy
Infoempresa.com	www.infoempresa.com	Spain
ISEA	www.iseamcc.net	Spain
Kenedict Innovation Analysis	www.kenedict.com	The Netherlands
Mozaika	www.mozajka.co	Bulgaria
Netage	netage.nl/en	The Netherlands
Normative	www.normative.io	Sweden
OECON GROUP	www.oecon.gr/en	Greece
Openlaws	openlaws.com/	Austria
OpenMove	openmove.com	Italy
Open Opps	www.openopps.com	United Kingdom
Peacemeetings	menplanner.org	Spain
SmartAppCity	www.smartappcity.com	Spain
Smartbow	www.smartbow.at	Austria
Smartup Cities	www.smartupcities.com	Spain
UniGraph	unigraph.rocks	United Kingdom
Wikimedia	www.wikimedia.org	Belgium
Wise Town	wise.town	Italy
Wripl technologies	wripl.com	Ireland
Zimmerman & Zimmerman	www.zimmermanzimmerman.nl	The Netherlands

Figure 62: Interviewed organisations

3. IODC Temperature Check

To complement and cross-check the findings of the survey as explained in the previous section, a temperature check was conducted at the International Open Data Conference 2016 (IODC) held in Madrid, Spain, on 6-7 October 2016. The IODC 2016 aimed to build stronger relationships between Open Data initiatives from the different governments and establish a dialogue between the different stakeholders. Titled 'Global goals, local impact' the fourth IODC focused on showcasing successes,

confront shared challenges, and help ensure that the Open Data vision and diverse Open Data initiatives continue to coordinate effectively[31].

The purpose of the temperature check was to collect information on how participants either used or perceived the use of Open Data. Obtaining a relevant sample thereby offers the opportunity to further understand the perception around the re-use of Open Data. A total of 92 stakeholders were asked to fill in a set of 6 questions relating to their experiences with Open Data. The first question related to the type of organisation the respondent represented. The following questions - relating to the main benefits, main challenges, type of data user, and lastly which domains were mainly used – were grouped based on the type of organisation the respondents represented. Since most organisations highlighted in this report come from the private sector, grouping the answers of the temperature check by type of organisation would provide the best opportunity to cross-check the findings of the survey, as shown in the figure below. The sixth question referred to the country the respective respondent was from. Respondents could choose more than one answer for every question. Especially regarding the main benefits, main challenges and main domains used, the majority of respondents provided more than one answer. Final percentages to each answer are based on the total amount of answers given per question. An analysis of each question is provided throughout the report.

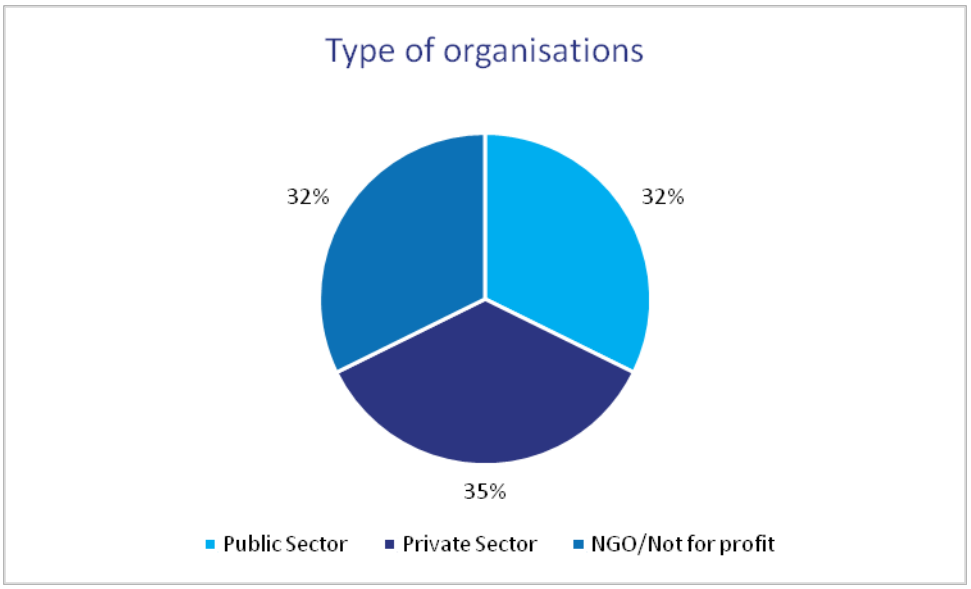


Figure 63: Type of organisations – IODC Temperature Check

[31] IODC 2016, Spain, Madrid

Annex II: List of organisations showcased

AppChallenge

AppChallenge is a British company providing open innovation competitions to clients with data who wish to engage with the 10,000+ app developers who belong to the AppChallenge community.
www.appchallenge.net

BBVA Data & Analytics

BBVA Data & Analytics is a consultancy firm on technology and analytics which is part of the global financial group BBVA. They provide organisations with competitive and sustained advantage through data and analytic services. BBVA Data & Analytics delivers analyses and provide raw aggregated data through an Open Data platform.
http://bbvadata.com/

Canard Drones

Canard Drones is a Spanish company providing services to calibrate the navigational aids at airports using unmanned aerial vehicles (drones) all over the world.
www.canarddrones.com/

Connemara Programme

The Connemara Programme is assisting businesses in Connemara, Ireland to drive a local recovery using local resources. All products and services are provided free to any community group or non-profit organisation based in a deployed area.
www.connemaraprogramme.com/

Crige PACA

Crige-PACA supports local authorities of the French region “Provence-Alpes Côte d’Azur” with the centralized collection and provision of geographical data.
www.crige-paca.org

CropDiagnosis

CropDiagnosis is a Greek mobile application which aims at improving pest management decisions of farmers. They do so by helping the farmer’s crop diagnosis becoming more accurate, selection of chemicals error-free and guidance on the application of the products.
www.cropdiagnosis.com

Doctrine.fr

Doctrine has a freemium business model. They offer a basic version of their services for free and more advanced services, such as legal analytics, for a fee. All users have access to the database with enriched, enhanced and linked Open Data on court decisions.
www.doctrine.fr

Egovlab

eGovlab is funded by the Swedish government, and delivers projects across a range of sectors from decision support systems and geospatial information systems to solutions aiming at mobile inclusiveness, democracy and smart communities.
www.egovlab.eu

eVineyard

eVineyard is a Slovenian application which analyses aggregated (Open) Data to help growers with the management of their vineyard. By combining multiple sources of data, this application helps the grower making irrigation, harvesting or treatment decisions in the vineyard.
www.evineyardapp.com

Fluicity

Fluicity provides solutions to face current problems in the field of citizen engagement. They have developed a platform which brings transparency and facilitates collaboration between citizens, the public space and the local government.
www.flui.city

Geomanagement

Geomanagement offers solutions based on data analyses. It uses open source software and Open Data to develop new applications and solutions.
http://levadia.gr

Glimworm Information Technologies

Glimworm Information Technologies is a technology company from the Netherlands which is active in the development of mobile applications, Internet of Things, beacons producer and many more technological fields. These elements come together in their Internet of Things Living Lab where they explore technological developments such as Open Data to experiment with IoT in Amsterdam.
www.glimworm.com

GraphDefined GmbH

GraphDefined assists companies in reaping the benefits of the data revolution. They offer consultancy services and software as a service.
www.graphdefined.de

Hammer Project

The Hammer Project developed an algorithm to extract the relevant data from Open Data.
www.hammer-project.com

Infoempresa.com

Infoempresa.com is an online provider of company information in Spain. They offer information from official sources which is updated daily (directly connected to Commercial Registry of Spain and other sources).
www.infoempresa.com

ISEA

ISEA S. COOP. is a private and non-profit innovation and entrepreneurship Centre, specialized in Business Services Sector, promoted by the Division of Engineering, and Business Services of MONDRAGON Corporation. ISEA uses Open Data in the TENTU project.
www.iseamcc.net

Kenedict Innovation Analysis

Kenedict enables organisations focused on innovation to significantly improve their market intelligence through a combination of network analytical services, advisory and training.
www.kenedict.com

Mozaika

Mozaika provides research and solutions in the field of data science, natural language semantics, discourse, and natural human computer interfaces, building information infrastructures to be used in data as a service (DaaS) contexts in a variety of applications including but not limited to creativity and research activity enhancement.
www.mozajka.co

Netage

Netage helps fire departments to get a grip on Open Data in their operational processes; either in the preparation phase or in the emergency response phase.
www.netage.nl

Normative

Normative provides a subscription based SaaS B2B model for companies to produce sustainability reports. They have developed an innovative tool, which uses machine learning to analyse the financial accounting of a company.
www.normative.io

OECON GROUP

The OECON group is a Greece-based management consultancy firm. They use Open Data for their studies, and they have developed a freely accessible data platform with Open Data.
www.oecon.gr/en

Openlaws

openlaws is a next generation legal information system. The platform automatically collects data from different sources (big data) and provides users with tools to increase productivity. The system helps small and medium enterprises (SMEs) in complying with legal requirements and supports large enterprises and corporations in meeting compliance requirements.
www.openlaws.com

OpenMove

OpenMove is an Italian mobile ticketing solution for the transport sector providing a platform which can be used both by the commuters and transportation agencies. This functionality makes it easier for the commuter to acquire tickets and provides a new sales channel for the providing agency.
www.openmove.com

OpenOpps.com

Open Opps uses Open Data to give insight into public procurement worldwide. It publishes tenders from across the globe openly, so that clients can access a world of opportunities.
www.openopps.com

Peace Meetings

Based on the collection of large amounts of data and social platforms, Peace Meetings offers peaceful solutions to social issues.
http://menplanner.org and http://peacemeetings.com

SmartAppCity

SmartAppCity brings all the available data and information together into one application, and creates a comprehensive application which is interesting for both visitors and citizens of a city to use.
www.smartappcity.com

SmartBow

SmartBow makes the ear tags of dairy cattle intelligent, thereby enabling farmers to monitor their cattle's behaviour and developments such as diseases, heat detection or the real time location of the individual cows.
www.smartbow.at

SmartUp Cities

SmartUp Cities develops technologies for smart city solutions, across a wide number of sectors. They transform data into actionable intelligence.
www.smartupcities.com

UniGraph

Unigraph provides access to unified and interconnected information coming from multiple Open Data sources to empower an ecosystem for data-driven decisions. They offer clients cleaned, organized and real-time data.
http://unigraph.rocks

Wikimedia

Wikimedia is a global movement whose mission is to bring free educational content to the world. Wikimedia strives to bring about a world in which every single human being can freely share in the sum of all knowledge.
www.wikimedia.org

Wise Town

WiseTown is an Italian application suite for the communication management between citizens and local governments on subjects such as urban quality, safety and shared planning while integrating social networks, Open data and IoT. The engine behind this application analyses, ranks and aggregates the data before assigning it to the responsible public body.
http://wise.town

Wripl Technologies

Wripl is online personalisation solutions provider that allows online businesses to instantly personalise their offering. For this wripl solutions provides high-end user engagement technologies to allow businesses to offer the right content to the right users at the right time and across devices.
http://wripl.com

Zimmerman & Zimmerman

Zimmerman & Zimmerman is a social enterprise which creates and develops solutions for organisations that need to become transparent about their financial project data.
www.zimmermanzimmerman.nl

Annex III: Correlation Matrix data categories

Correlation matrix of the re-use of Open Data categories

	Agriculture, Fisheries, Forestry & Foods	Energy	Regions & Cities	Transport	Economy & Finance	International Issues	Government & Public Sector	Justice, Legal System & Public Safety	Environment	Education, Culture & Sport	Health	Population & Society	Science & Technology	Other, namely..
Agriculture, Fisheries, Forestry & Foods	1.00													
Energy	0.42	1.00												
Regions & Cities	0.36	0.18	1.00											
Transport	0.33	0.26	0.60	1.00										
Economy & Finance	0.10	0.24	0.24	0.22	1.00									
International Issues	0.23	0.22	0.19	0.18	0.28	1.00								
Government & Public Sector	0.09	0.23	0.11	0.09	0.24	0.19	1.00							
Justice, Legal System & Public Safety	0.10	0.09	0.05	0.12	0.12	0.35	0.31	1.00						
Environment	0.50	0.31	0.52	0.42	0.14	0.17	0.23	0.10	1.00					
Education, Culture & Sport	0.20	0.24	0.24	0.23	0.04	0.22	0.02	0.16	0.27	1.00				
Health	0.49	0.34	0.26	0.40	0.12	0.35	0.04	0.21	0.31	0.44	1.00			
Population & Society	0.22	0.21	0.68	0.48	0.35	0.22	0.22	0.08	0.51	0.23	0.28	1.00		
Science & Technology	0.04	0.34	0.13	0.17	-0.02	0.33	0.02	0.04	0.09	0.17	0.12	0.11	1.00	
Other, namely..	0.16	0.21	0.03	0.07	0.09	0.27	0.08	0.24	0.06	0.20	0.24	0.05	0.22	1.00

Correlations higher than 0.5 yet smaller than 1 are marked in green

Correlations lower than 0.1 are marked in orange

Figure 64: Correlation matrix data categories

Annex IV: Interview questionnaire

[Company]

Date:

Interviewee:

Interviewer(s):

[URL]

Getting to know each other

Brief explanation of this research and purpose of the interview

1. Providing the interviewee the opportunity to introduce himself

Company Summary

Discussion about the history, organisation and works of Fluicity

2. How would you describe your business model in one sentence?

Open Data, for you and your clients

3. What are the main benefits you see in making use of Open Data?
4. Do you communicate on Open Data with your customers and how do you understand their perception?

Working with Open Data

5. How would you define open data?
6. Could you describe the typical profile or key characteristics of the skills needed for someone working with data? Are there any skills needed in particular to work with Open Data?
7. Do you provide training or other forms of education to your employees for them in order to obtain these skills?
8. How do you foresee the future for recruiting new employees with these skills?

The challenges with Open Data

9. What are the main barriers faced when working with Open Data?
10. When looking back at these barriers, what would you have done differently?
11. Do you see technical aspects as being the main bottleneck of working with Open Data?
12. What would be a success factor to work with Open Data?
13. If you had a wish list for specific datasets to be made available, what would these be?
14. What should be done to improve access to data?

Room for discussion

15. Concluding comments on the re-use of Open Data by the interviewee

Other points

Questions & comments discovered during the interview

Annex V: Survey questionnaire

Survey questionnaire on the benefits of re-using Open Data

The definition of Open Data used in the survey is:

Open (Government) Data refers to the information collected, produced or paid for by the public bodies (also referred to as Public Sector Information) and made freely available for re-use for any purpose.

Please answer all questions to the best of your knowledge.

Questionnaire

General

Note: To check the box of the correct answer, please double click on the grey box.

1.1 Company Name: _____

1.2 Company URL: _____

1.3 City: _____

1.4 Country: _____

1.5 Other locations of the company, including main locations of your users/clients

1.6 Founding year:_____

1.7 Who owns the company?
☐ Shareholders
☐ Family owned
☐ Two business owners
☐ The CEO
☐ The management team (more than 2)

1.8 Is your company a:
☐ For-profit organisation
☐ Non-profit organisation

1.9 Who are you in business with?
☐ Business to Business

☐ Business to Customer
☐ Business to Government
☐ None of the above

1.10 Sector (NACE code)

☐ Agriculture, forestry and fishing
☐ Mining and quarrying
☐ Manufacturing
☐ Electricity, gas, steam and air conditioning supply
☐ Water supply, sewerage, waste management and remediation activities
☐ Construction
☐ Wholesale and retail trade, repair of motor vehicles and motorcycles
☐ Transportation and storage
☐ Accommodation and food service activities
☐ Information and communication
☐ Financial and insurance activities
☐ Real estate activities
☐ Professional, scientific and technical activities
☐ Administrative and support service activities
☐ Public administration and defence, compulsory social security
☐ Education
☐ Human health and social work activities
☐ Arts, entertainment and recreation
☐ Other service activities
☐ Activities of households as employers, undifferentiated goods- and services-producing activities of households for own use
☐ Activities of extra territorial organisations and bodies

1.11 Significant sources of revenue for your company?

☐ Selling products (applications, physical products)
☐ Selling services
☐ Advertisement
☐ Other, namely...
☐ if other sources of revenue

1.12 What would be the **most** significant source of revenue for your company?

1.13 What is your business model around the use of Open Data?

☐ Enhancing products
☐ Enhancing services
☐ Process optimisation
☐ Data as a service
☐ Information as a service
☐ Answers as a service
☐ Development of web or mobile applications

1.14 Which business model describes your company best?

☐ Data Supplier
☐ Data Aggregator
☐ Data Developer
☐ Data Enricher
☐ Data Enabler

Archetype	Description
Supplier	Allows others to use and re-use Open Data
Aggregator	Collects and aggregates Open Data and, sometimes, other proprietary data and finds correlations, identifies efficiencies, and visualizes complex relationships
Developer	Designs, builds, and[/or] sells web-based, tablet, smartphone applications
Enricher	Uses Open Data to enhance their existing products and services through better insight (typically larger companies)
Enabler	Facilitates the supply of Open Data or use of Open Data

1.15 What are your distribution mechanisms?

☐ Mobile applications via application store
☐ Webshop or website
☐ Office or store
☐ If other distribution, namely

1.16 Short description of the company's activities, including the main products and services developed:

Financial Information

2.1 Average annual turnover for the past 5 years (or if the company is < 5 years, since the founding)?

- ☐ 0 – 9,999 euro
- ☐ 10,000 – 49,999 euro
- ☐ 50,000 – 99,999 euro
- ☐ 100,000 – 249,999 euro
- ☐ 250,000 – 499,999 euro
- ☐ 500,000 – 999,999 euro
- ☐ 1,000,000 – 4,999,999 euro
- ☐ > 5,000,000 euro

2.2 What was your turnover in 2015 or 2014, if the figures for 2015 are not available yet?

- ☐ 0 – 9,999 euro
- ☐ 10,000 – 49,999 euro
- ☐ 50,000 – 99,999 euro
- ☐ 100,000 – 249,999 euro
- ☐ 250,000 – 499,999 euro
- ☐ 500,000 – 999,999 euro
- ☐ 1,000,000 – 4,999,999 euro
- ☐ > 5,000,000 euro

2.3 Which percentage of your 2015 turnover is linked **directly** to the re-use of Open Data resources (Open Data may be one of the several inputs)? Please try to give your best estimate.

- ☐ 0 – 5%
- ☐ 6 – 10%
- ☐ 11 – 20%
- ☐ 21 – 30%
- ☐ 31 – 40%
- ☐ 41 – 50%
- ☐ 51 – 60%
- ☐ 61 – 70%
- ☐ 71 – 80%
- ☐ 81 – 90%
- ☐ 91 – 95%
- ☐ 96 – 100%

2.4 Which percentage of your 2015 turnover is linked to **indirect** benefits of the re-use of Open Data resources: efficiency gains, production gains, time gains (Open Data may be one of the several inputs)? Please try to give your best estimate.

- ☐ 0 – 5%
- ☐ 6 – 10%
- ☐ 11 – 20%
- ☐ 21 – 30%
- ☐ 31 – 40%
- ☐ 41 – 50%
- ☐ 51 – 60%
- ☐ 61 – 70%
- ☐ 71 – 80%
- ☐ 81 – 90%

☐ 91 – 95%
☐ 96 – 100%

2.5 Forecasted annual growth in general turnover for the upcoming 2 years?

☐ 0 – 5%
☐ 6 – 10%
☐ 11 – 20%
☐ 21 – 30%
☐ 31 – 40%
☐ 41 – 50%
☐ 51 – 60%
☐ 61 – 70%
☐ 71 – 80%
☐ 81 – 90%
☐ 91 – 95%
☐ 96 – 100%
☐ 101 – 110%
☐ 111 – 120%
☐ 121 – 140%
☐ 141 – 160%
☐ 161 - 180%
☐ 181 - 200%
☐ 201 - 250%
☐ 251 - 300%
☐ >300%

2.6 Forecasted annual growth in turnover **related to Open Data** for the upcoming 2 years?

☐ 0 – 5%
☐ 6 – 10%
☐ 11 – 20%
☐ 21 – 30%
☐ 31 – 40%
☐ 41 – 50%
☐ 51 – 60%
☐ 61 – 70%
☐ 71 – 80%
☐ 81 – 90%
☐ 91 – 95%
☐ 96 – 100%
☐ 101 – 110%
☐ 111 – 120%
☐ 121 – 140%
☐ 141 – 160%
☐ 161 - 180%
☐ 181 - 200%
☐ 201 - 250%
☐ 251 - 300%
☐ >300%

2.7 Please include any financial or operational information that can help further understand your company and its activities (can be attached to the email)

People

3.1 Number of **employees** in FTE's (the amount of employees working at the company converted to the fulltime equivalent).

- ☐ 0 – 5
- ☐ 6 – 10
- ☐ 11 – 15
- ☐ 16 – 25
- ☐ 26 – 50
- ☐ 51 – 100
- ☐ 101 – 250
- ☐ 251 – 1000
- ☐ >1000

3.2 Number of **externals** in FTE's (externals are people that are working for the company on project basis, but are not on the company's payroll)

- ☐ 0 – 5
- ☐ 6 – 10
- ☐ 11 – 15
- ☐ 16 – 25
- ☐ 26 – 50
- ☐ 51 – 100
- ☐ 101 – 250
- ☐ 251 – 1000
- ☐ >1000

3.3 Average number of employees working at the company over the past 5 years?

- ☐ 0 – 5
- ☐ 6 – 10
- ☐ 11 – 15
- ☐ 16 – 25
- ☐ 26 – 50
- ☐ 51 – 100
- ☐ 101 – 250
- ☐ 251 – 1000
- ☐ >1000

3.4 Average number of employees directly working on Open Data products or services (i.e. involved in working with data)?

- ☐ 0 – 5
- ☐ 6 – 10
- ☐ 11 – 15
- ☐ 16 – 25
- ☐ 26 – 50
- ☐ 51 – 100

☐ 101 – 250
☐ 251 – 1000
☐ >1000

3.5 Is your company foreseeing to employ further individuals?

☐ yes
☐ no

3.6 if yes to 3.5: Which profiles are you looking for predominantly?

☐ Marketing
☐ Data scientist
☐ Sales
☐ Content expert
☐ Manager
☐ Other, namely..

3.7 if yes to 3.5: Which skills are you looking for predominantly?

☐ Statistical skills (analysing data)
☐ Analytical skills (problem solving)
☐ Technical skills (programming, algorithms)
☐ Communication skills
☐ Business Insight (strategy, advisory)
☐ Domain Knowledge / work experience

3.8 How would you rank the skills mentioned below with **6** being the most important skills and **1** being the least important for your organisation?

☐ Statistical skills (analysing data)
☐ Analytical skills (problem solving)
☐ Technical skills (programming, algorithms)
☐ Communication skills
☐ Business Insight (strategy, advisory)
☐ Domain Knowledge / work experience

Use of Data Sets

4.1 Which of the 13 categories do you re-use data from?

☐ Agriculture, forestry and fishing
☐ Energy
☐ Regions and cities
☐ Transport
☐ Economy & Finance
☐ International Issues
☐ Government & Public Sector
☐ Justice, Legal System & Public Safety

- ☐ Environment
- ☐ Education, Culture & Sport
- ☐ Health
- ☐ Population & Society
- ☐ Science & Technology
- ☐ Other, namely..

4.2 Which of the 5 high priority data domains do you re-use data from?

- ☐ Statistics
- ☐ Geospatial
- ☐ Transport and infrastructure
- ☐ Companies
- ☐ Earth observation
- ☐ None of the above

4.3 From which of the countries below do you collect data?

4.4 From which regions do you use data?

- ☐ Africa
- ☐ Antarctica
- ☐ Asia
- ☐ Central America
- ☐ European non-EU country
- ☐ European Union
- ☐ Middle East
- ☐ North America
- ☐ Oceania
- ☐ South America
- ☐ The Caribbean
- ☐ None of the above

4.5 Platform by which the data was accessed?

- ☐ European Data Portal
- ☐ National Open Data Portal
- ☐ Regional or Local Open Data Portal
- ☐ Provided by the public organisation directly
- ☐ Other, namely..

4.6 Is it easy to find data you need?

- ☐ Very difficult

☐ Difficult
☐ Easy
☐ Very easy

4.7 On a scale of 1 to 4, how would you rate the usefulness of Open Data datasets? (1- poor, 4- excellent) Your answer can reflect your experience with data quality, format of the data, or other factors.

☐ Poor
☐ Average
☐ Good
☐ Excellent

4.8 Why did you give it this rating?

4.9 On a scale of 1 to 4, how would you rate the availability of Open Data datasets? (1- poor, 4- excellent)

☐ Poor
☐ Average
☐ Good
☐ Excellent

4.10 What datasets (if any) are not currently available that would be useful for your company to have as Open Data?

Success Factors

5.1 What are critical sources of data for your company?

☐ My company is relying on quality data
☐ My company is relying on real-time data
☐ My company is relying on data from a specific organisation
☐ My company is relying on a specific dataset
☐ My company is relying on the systematic and continued publication of specific data sets
☐ My company does not use data as a critical business source

5.2 Besides revenue generation, how do you measure the impact your company has on society and the public good?

☐ Social media monitoring to retrieve the public opinion
☐ Measuring impact via interviews or surveys
☐ Feedback received from clients

☐ I don't measure that

5.3 What is the most interesting impact your company established? Think about how it had an effect on behaviour, the environment, process optimisation, cost reduction etc.

5.4 Do you have a language preference in re-using data sets? If yes, which one?

5.5 Does language matter?

☐ No, strongly disagree
☐ No, disagree
☐ Yes, agree
☐ Yes, strongly agree

5.6 Why would you say language matters?

5.7 Do you collect feedback from users regarding the services and products you deliver to improve?

☐ Yes, regularly
☐ Yes, sometimes
☐ No

5.8 Do you collaborate with others?

☐ Yes, other individual companies
☐ Yes, I am part of a community
☐ Yes, the public sector
☐ Yes, during events
☐ No

Contact Information

6.1 First and last name:

_____

6.2 Job title:

_____

6.3 Email: _____

6.4 Phone: _____

Other

7.1 Recommendations: What other companies, either in your sector or other sectors, would you recommend we contact regarding their use of Open Data?

7.2 What conferences or events do you think would be helpful to us in surveying the field of Open Data companies?

Thank you for your kind cooperation.

Annex VI: Temperature check during IODC

1. What organisation do you come from?
    - Public Sector
    - Private Sector
    - Non-Governmental Organisation/ Not for profit
2. What are the main benefits you see in making use of Open Data?
    - The (limited) cost
    - The societal value
    - The scope of available information
    - Creating new products
    - Other;

3. What are or creates the main challenges in making use of Open Data?
    - The quality
    - The availability
    - The metadata
    - The format
    - The standardisation
    - Other;

4. Are you a:
    - Data Supplier
    - Data Aggregator
    - Data Developer
    - Data Enricher
    - Data Enabler
5. From which category does the data you use come from?
    - Agriculture, Fisheries, Forestry & Foods
    - Energy
    - Regions & Cities
    - Transport
    - Economy & Finance
    - International Issues
    - Government & Public Sector
    - Justice, Legal System & Public Safety
    - Environment
    - Education, Culture & Sport
    - Health
    - Population & Society
    - Science & Technology
    - Other, namely

Email address & URL: (optional)

Country: (optional)

Would you like to be contacted by us if we have further questions?