EUROPEAN
DATA PORTAL

The Economic Impact of Open Data
Opportunities for value creation in Europe

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 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:
Esther Huyer
Email: esther.huyer@capgemini.com
Laura van Knippenberg
Email: laura.van.knippenberg@capgemini.com

With contributions by:
Eline Lincklaen Arriëns
Marit Blank

Last update: 26.01.2020
www: https://europeandataportal.eu/
@: info@europeandataportal.eu

PDF ISBN: 978-92-78-42108-3 doi: 10.2830/63132 OA-04-20-043-EN-N

Executive summary

"Data and AI are the ingredients for innovation that can help us to find solutions to societal challenges, from health to farming, from security to manufacturing. In order to release that potential we have to find our European way, balancing the flow and wide use of data while preserving high privacy, security, safety and ethical standards."

European Commission President
Ursula von der Leyen in "A Union that strives for more - My agenda for Europe" (2019)[1]

The 7 key learnings

1. The specification and implementation of high-value datasets as part of the new Open Data Directive is a promising opportunity to address quality & quantity demands of open data.
2. Addressing quality & quantity demands is important, yet not enough to reach the full potential of open data.
3. Open data re-users have to be aware and capable of understanding and leveraging the potential.
4. Open data value creation is part of the wider challenge of skill and process transformation: a lengthy process whose change and impact are not always easy to observe and measure.
5. Sector-specific initiatives and collaboration in and across private and public sector foster value creation.
6. Combining open data with personal, shared, or crowdsourced data is vital for the realisation of further growth of the open data market.
7. For different challenges, we must explore and improve multiple approaches of data re-use that are ethical, sustainable, and fit-for-purpose.

[1] "A Union that strives for more - My agenda for Europe" (2019) available at: https://ec.europa.eu/commission/sites/beta-political/files/political-guidelines-next-commission_en.pdf

Executive summary

Open data market size

- €184.45 billion open data market size in 2019
- €199.51 - €334.20 billion open data market size forecast for 2025

Open data employment

- 1.09 million open data employees in 2019
- 1.12 – 1.97 million open data employees forecast for 2025

Open data potential per sector

- 15.7% growth expected from high impact and high potential sectors
- High impact:
- High potential:

For details on calculations and assumptions see corresponding sections.

EUROPEAN DATA PORTAL

Executive summary

Efficiency gains

- Saving lives, e.g. 54 - 202 thousand lives saved by faster emergency response
- Saving time, e.g. 27 million hours saved in public transport
- Saving the environment, e.g. 5.8 Mtoe* saved by reducing household energy consumption
- Improving language services with open data, e.g. by increasing machine translation

Cost savings

- Saving healthcare costs, e.g. €312 - €400 thousand due to faster first aid by bystanders
- Saving labour costs, e.g. €13.7 - €20 billion by reducing time spent in traffic
- Saving costs on energy bills, e.g. €79.6 billion due to more solar energy production
- Saving public sector costs, e.g. €1.1 billion by lower translation costs

Open data in organisations

- 49% of data used by surveyed organisations is open data and 77% of organisations plan to use more data
- 46% of organisations' revenues are impacted by open data and 73% of organisations expect the impact to increase
- 70% of surveyed organisations create data internally, of which 58% publish some of it as open data

* Million tonnes of oil equivalent

For details on calculations and assumptions see corresponding sections.

EUROPEAN DATA PORTAL

LIST OF FIGURES

LIST OF TABLES

GRAPHICAL ELEMENTS

Icons made by Freepik from www.flaticon.com available at: https://www.flaticon.com/authors/freepik

Chapter overview

TABLE OF CONTENTS

1. INTRODUCTION

1.1 Understanding the value of open data by exploring its economic impact

Since the start of the open data movement, thousands of public datasets were opened across Europe enabling new applications and insights. Moreover, it started a switch of mindset about public data as a common good - similar to infrastructure - to enable a fair and thriving European economy. Open data became a vital part of the European digital agenda and Member States implemented their national strategies accordingly.

The economic impact that was expected to be high and very tangible turned out to be more incremental and not immediately visible because it is subtle and sometimes well hidden. The benefits of open data are created by insights that enrich research and inform decision-making, by services in the form of apps and websites, by improved products and processes that increase productivity and efficiency, wellbeing, health, safety, and sustainability. Quantifying the benefits to express the economic impact of open data is comparably complex because the most important and significant benefits are indirect.

Similar to infrastructure, open data is an **enabler** for the economy. As an enabler, its impact cannot be measured by simply taking into account the costs for - to stay with the infrastructure analogy - construction of roads, canals, airports, train stations, and revenues created by tolls. The value lies in the fact that people, organisations, and products are enabled to move from A to B as well as in people, organisations, and products that are needed and used to do so. Even if one does not know how to sail a big ship and has no intention to ever sail on one, she may still benefit from the goods shipped on these ships. She surely benefits from the employment generated by the cargo shipping company, the tax they pay and that returns to her translated into public spending and economic stability.

For the realisation and maintenance of a new infrastructure project, for example in energy and water supply, the social, economic, environmental and political risks, costs, and benefits are explored, quantified, and weighted. The same logic applies to the realisation and maintenance of open data projects: the benefits have to be explored and quantified to give a reasonable insight into its value compared to costs of resources. Furthermore, it increases our understanding of where it is most beneficial to allocate resources. If we understand the economic benefit, we can keep open data high on the political agenda. It helps to communicate more clearly that making data open is not an altruistic act of charity but a vital contribution to enable a growing, innovative, ethical economy as a vital pillar for, democracy, safety and welfare.

Similar to looking at our transport network in a holistic way when planning how to move from Prague to Dublin, this report looks at data from a holistic perspective. Just like a road to a place in no-man's-land can seem of little value, if at the end of that road a harbour or an airport is located, this road is crucial. In the same way, open data on diagnostics and health statistics linked to a personal health record can be crucial to understand symptoms. Therefore, this report looks at the economic impact of open data as part of the data economy as a whole. We will shed light on the importance of open data as a catalyser for the EU data economy and the European economy as a whole.

We bring our results into context to proof their feasibility. We will critically look back and consider what the open data community observed in learned in the last ten to 15 years. We reflect on what that means for the present day and for the future outlook.

1.1.1. Report structure

1. Introduction
2. Open data market size
   The report starts at the European macro-economic level exploring aspect of the market for open data namely the open data market size and potential market growth. In addition, the results are put into context to understand their implications better.
3. Open data employment
   The next chapter focuses on jobs in the open data market looking at direct impact on jobs working with or due to open data.
4. Open data potential per sector
   After understanding the overall market size and job market better, the report explores the economic potential of open data in different sectors and forecasts the open data gain per sector for 2025.
5. Efficiency gains due to open data
   In the next part, the report looks at specific impact cases in more detail to enrich, validate and contextualise the metrics with case studies and real-life examples to show how value is created at a very practical level. Causal value chains and calculations are explained to show the level of direct and indirect impact. The impact of open data on saving lives and saving time is shown, environmental benefits and equal access to knowledge and technology is explored.
6. Saving costs due to open data
   The following chapter looks at different angles of savings costs due to open data namely the reduction in cost spend on data and secondly the cost saved by processes being more efficient and productive.
7. Open data in organisations
   This chapter introduces the report's findings of a survey with European companies self-assessing the economic impact of open data in their organisations. Of these companies, 28 case studies can be found in de appendix in which they tell their story of how impact is created with open data.
8. Discussion

1.2 Lessons learned from relevant literature

The value of open data by the European Data Portal in 2015

In 2015, the European Data Portal published its first report on the value of open data in Europe (EU28+)[2]. Since then, more relevant literature was conducted and made available. This enables the research community to compare and refine their findings and update their methodology. To understand how and why this report follows or diverges from the previous methodology, the approach in 2015 is explained and, in addition to that, lessons learned from other studies in the field are taken into account.

The 2015 study "Creating value through open data" focused on the following indicators:

- Open data market size: the direct and indirect open data market size is calculated based on MEPSIR (2006)[3] data on the open data market as % of GDP. To forecast the open data market, real GDP forecasts and the open data maturity classification of 2015 are used. The open data market is estimated at €193 - €209 billion in 2016 and forecast to be €265 – €286 billion in 2020. Additionally, the distribution of the

[2] European Data Portal (2015) available at: https://www.europeandataportal.eu/sites/default/files/edp_creating_value_through_open_data_0.pdf
[3] MEPSIR (2006) available at: http://ec.europa.eu/newsroom/document.cfm?doc_id=1197

open data market across sectors is calculated using Eurostat figures on GDP per sector combined with a research by DemosEuropa & WISE[4].

- Open data jobs: the direct and indirect open data jobs are measured as the number of persons employed. In the 2015 report, the difference between direct and indirect is the following: *"Start-ups that are using open data for their business are directly creating open data jobs, like open data analysts or product marketers. If these businesses that make use of open data also need a sales manager for their company as it gets bigger, this would be an indirect job that is created"* (EDP 2015, p.66). To estimate the number of direct open data employees, a series of Spanish studies from Asedie[5] researching the number of people directly employed in the Spanish private open data sector (thus not including public sector) are used. To forecast the number of open data employees, the Eurostat employment forecast and Cedefop (2010)[6] skills forecast are used. The total (direct and indirect) open data job market is estimated at 235 700 - 284 900 in 2016 and forecast to be 312 600 - 377 000 in 2020.

- Cost saving in public sector: the cost savings in the public sector are calculated based on the expected cost savings for the Danish government due to open data. This is expressed as a percentage of the total public sector expenditures for Denmark and subsequently pro-rated for Europe. The accumulated cost savings for 2020 are forecasted to equal €1.7 billion.

- Efficiency gains: several efficiency gains are qualitatively described through examples and where possible quantification of these gains is provided, such as:
  - 7 000 potential lives saved due to receiving CPR from bystanders in case of cardiac arrest
  - 1 425 potential lives saved due decrease of road fatalities
  - 629 million hours saved due to less time spent stuck in traffic

The methodology used in this report is for a large part along the lines of the previous report to make the results comparable. For example, this report - similar to the 2015 report - considers the open data market size and employment to be linked to GDP and employment rates. However, this report enhances the methods in several ways, in order to leverage new insights and more recent numbers that have become available. For example, growth rates in the 2020 study are informed by more recent literature results, include more recent open data maturity figures (2019)[7], and consider the maturity trend. This provides the opportunity to reflect on the insights derived in 2015, yet, due to the slight differences in approach, the results need to be contextualised when comparing them. In addition to methodology adaptations, another difference is the switch from performing calculations on the EU28+ countries in 2015 to the focus of this report on the EU27+ countries. This again affects in parts the comparability. Besides critically assessing the methodology of the 2015 report, there is more available literature from which valuable insights can be derived to optimise the method of this report.

Over the past years, several studies assessing the economic impact of open data have been published on both country specific as well as on international level. The Kyiv School of Economics, for example, used an ex-ante approach to estimate the impact of open data on the Ukrainian economy (2018)[8]. The Luxembourg Institute of Science and Technology also used ex-ante indicators to forecast the open data

4 DemosEuropa and Warsaw Institute for Economic Studies (2014) available at: https://tech.eu/features/381/open-big-data-in-europe/
5 ASEDIE (2019) available at: http://www.asedie.es/assets/asedie.-infomediary-sector-report-2019.pdf
6 Cedefop (2010) available at: http://www.cedefop.europa.eu/ files/3052_en.pdf
7 European Data Portal (2019) available at: https://www.europeandataportal.eu/sites/default/files/open_data_maturity_report_2019.pdf
8 The Kyiv School of Economics (2018) available at: http://tapas.org.ua/en/media/ekonomichnyj-potentsial-vidkrytykh-danykh-v-ukraini/

market size, the number of jobs, and the cost saved by open data in Luxembourg (2018)[9]. The study also used a bottom-up approach by a questionnaire aimed at open data re-users in the private and public sector. An example of an international study is the Socio Economic Impact of Open ELS (2018)[10], which - through a combination of desk research, interviews, and a survey targeting SMEs - highlights the benefits of open geospatial data on the overall geospatial market at the European level.

In our literature research, we focused on the most recent studies in the open data field, i.e. studies that have been published between 2017 and 2019. Besides the most recent studies, we also reviewed the most influential studies on the economic impact of open data published before 2017. In addition, we included studies on the economic impact of data, since data and open data are closely linked to each other, and open data should therefore not be examined and analysed as an isolated concept.

Studies that have been conducted in previous years apply different methodologies and approaches to measure the economic impact of open data. Several lessons can be learned from these approaches. The key lessons learned from our literature research are summarised below:

- **Applied methodologies are not transparent.** Studies into open data apply different methodologies to measure impact. These methodologies are often not well described if described at all. Some studies include a methodology chapter, but even in these cases, the level of detail is very modest. It is often unclear how results are calculated, and which numbers and resources are used in these calculations. The lack of transparency makes it hard to verify results and replicate the studies.

- **Assumptions are not made explicit.** In line with the lack of transparency in methodologies, studies are often not transparent in their underlying assumptions. Findings are extrapolated, without providing the necessary arguments and without providing insights into the underlying assumptions. This makes it hard to understand the thoughts behind applied approaches and to understand how robust the results are.

- **Unclear link between literature and methodology.** It is often unclear if the applied approaches and methodologies are based on existing literature. Although some studies provide extensive literature overviews, the link between literature and methodology is not explicit and thereby often unclear.

- **The scope of open data is often unclear.** In line with the above-mentioned different taxonomies, it can be observed that the concept of open data is, in some cases, used in a very broad way without indicating which definition of open data is used in the study. Concepts as data, open data, public sector information, and open government data, are used as interchangeable concepts. Moreover, some studies calculate the impact of opening any kind of data, even personal data. This is, of course, possible under certain circumstances and hence the potential impact not wrong just highly optimistic and neglecting several legal, ethical and technical considerations that would need to be solved first.

- **Either macro-economic or micro-economic approaches.** Studies often focus solely on several different macro-economic indicators using quantitative analyses, or solely on micro-economic indicators using qualitative analysis. Few studies aggregate both approaches and combine quantitative, qualitative, and case-study analyses.

[9] The Luxembourg Institute of Science and Technology (2018) available at: https://download.data.public.lu/resources/study-impacts-of-open-data-in-luxembourg-and-the-greater-region-2018/20181004-093205/impacts-of-open-data-in-luxembourg-and-the-greater-region-2018.pdf

[10] Socio Economic Impact of Open ELS (2018) available at: https://openels.eu/wp-content/uploads/2019/04/Open_ELS_socio_economic_benefits_final_report_Website.pdf

- **Reliance on selective number of secondary sources.** Few studies gather primary data, for example by using questionnaires, conducting interviews, or expert discussions. Studies often use solely secondary data for their research and this secondary data comes from a selective number of studies that are relatively outdated.

- **Studies are difficult to compare.** Studies have different scopes and use different taxonomies, which makes it hard to compare them. Some studies focus on specific sectors, others on specific countries, or on specific types of data, such as geographical data. Differences in scope are not necessarily problematic when the same taxonomy is used. However, this is often not the case. Different definitions are used, for example, different sector definitions or open data definitions.

Table 1 Comparing terminologies

Comparing terminologies: data economy vs. impact of open data	
Terminology in the European Data Market Study by IDC (2017)[11]	**Terminology in this study**
Data workers: "Data workers are defined as workers who collect, store, manage and analyse data as their primary, or as a relevant part of their activity. (...) They elaborate and visualize structured and unstructured data to support analysis and decision-making processes."	Open data employees[12]: People (not FTE) in both the public and the private sector, that are generating, providing, aggregating, re-using, and enriching[13] open data or substantially enable others to do so in organisations that base their business case or core focus on open data. People (not FTE) in both the public and the private sector, that are generating, providing, aggregating, re-using, and enriching open data in organisations that do not base their business case or core focus on open data.
Employment share: Employment share is given by the share of data workers on the total employment in Europe, in percentage.	Employment share: Employment share is the share of direct and indirect data employees on the total employment per European country, for EU27+ and 28+, in percentage.

11 IDC (2017) available at: https://a2528ba5-a-c3c32646-s-sites.googlegroups.com/a/open-evidence.com/download/repository/SMART20130063_Final%20Report_030417_2.pdf?attachauth=ANoY7crL55TcXaDjVhYlzdD2jBNz6-UaN35LRS8sfZ9bqquqBazhj8Xk5FFQaus6z1xdbaviMZQ0Wbjkhst09Wpv7a13pztoDD9FrCANHECqEHeODrZgkvAvEes5w0ZiwC5t8L0Ovpk_7CHJVVPmUZWeUiVCYQoRGYk_b5ln0ZeTP3TH_A2sJhNeFtAg8X0ap3CqXfqHFmJGtgZkkEr0bcyM0pbRtH7UZ76oxlk89xXYB3FhdJ096eQqfhBG6Cx0TBGI9VCSKQRE&attredirects=0; http://datalandscape.eu/webinars/how-power-data-can-drive-european-union-economy-european-data-market-study-and-monitoring

12 European Data Portal (2015) used open data jobs. This indicator was renamed in the 2019 study to be distinguish beytween employees and FTE.

13 Based on the Data Value Chain Archetypes of the European Data Portal (2015, p.29) available at: https://www.europeandataportal.eu/sites/default/files/edp_creating_value_through_open_data_0.pdf

1.3 Our approach of exploring the economic impact of open data

1.3.1 Methodology overview

Based on the above-mentioned lessons learned from previous research and based on our experience from measuring the economic impact of open data in 2015[14], we have designed the methodological approach. To shed as much light as possible on the complex field of the economic impact of open data, the report combines a mixed approach of micro and macro-economic observations and calculations using primary and secondary research.

Macro-economic estimations entail quantitative analyses based on Gross Domestic Product (GDP) estimates and similar national views, whereas **micro-economic estimations** provide quantitative analyses based on company- or individual-level data. **Case study-**based approaches deploy qualitative or mixed methods based on specific domains with a selection of cases.

The exact methodology for specific chapters and calculations are provided in each chapter, as they are complex and best understood when applied and linked to the results. Throughout the report, when explaining the methodology, further references to methodologies from EDP (2015), the European Data Market Study (2017) or other relevant studies are made.

The table below shows an overview of the methodological approaches and selected indicators.

Table 2: Overview of methods in each chapter

Market size	Calculations, extrapolation, top down and bottom-up, quantitative	Primary and secondary data
Job market	Calculations, extrapolation bottom-up, quantitative	Primary and secondary data
Sectors	Calculations, extrapolation, examples, top down and bottom-up, quantitative	Primary and secondary data
Efficiency	Calculations, extrapolation, examples and causal value chains, bottom-up, quantitative and qualitative	Secondary data
Costs saved	Calculations, extrapolation, examples and causal value chains, bottom-up, quantitative and qualitative	Secondary data
Organisations	Case studies and survey, qualitative	Primary and secondary data

[14] European Data Portal (2015) available at: https://www.europeandataportal.eu/sites/default/files/edp_creating_value_through_open_data_0.pdf

1.3.2 The scope of data and impact

This report focuses on the economic impact of re-using open data. Open data includes open government data and open data from the private sector. In addition, we are including examples that show how open data is combined with shared data or personal data to highlight the fact that open data has significant potential when combined with other data and that looking at the data economy holistically is the most fruitful and solution-orientated approach.

Impact levels from immediate to complicity

The report, as it is common practice, distinguishes between different layers of impact depending on how immediate impact is created. Each example might have different levels of impact and therefore, the degree to how immediate the impact is always indicated. Impact can range between:

- Direct impact: clear and immediate relation between the factors and the impact.
- Indirect impact: the relation between factors and impact can be shown using examples and arguments. Generally, more assumptions are used to derive statements.

Status, potential, and forecast

Open data impact can already be realised, or it can be a future potential. Instead of the current status, often, open data impact is indicated as a potential that can be reached if we would make more use of certain datasets and applications. This report, where possible, provides both current status and a forecast. When talking about the future we distinguish between a forecast and an evaluation of the potential. A potential does not include the likelihood or timespan of its realisation. A forecast makes an assumption about the time span in which a potential can be realised. The actual level of realisation differs between countries, regions, sectors, etc.

The geographical scope and extrapolation

This report focuses on Europe. Therefore, macro-economic calculations look at Europe as a whole and whenever possible also provide national results. The causal value chains are extrapolated to Europe based on, for example, inhabitants or cities wherever it makes sense and adds insight. In many cases, local and regional differences and individual characteristics are not taken into account to allow for European-wide statements. Those cannot always be applied to individuals or specific situation without contextualisation.

1.3.3 Our credo: accuracy, transparency, modesty, and curiosity

We aim at achieving accuracy in our methods. Where precision is not feasible or not reasonable, we are transparent about this fact. Calculations are fully transparent, and either explained in the text or added in the footnotes. This allows the reader to understand the thoughts behind the calculations, challenge, update or further develop them. It also allows to use similar formulas but combine them with local or national data instead of European-wide data to derive insights about a region or country. We explore the market with modesty and curiosity. Results are derived in the ambition to understand the economic impact of open data better and not to promote it.

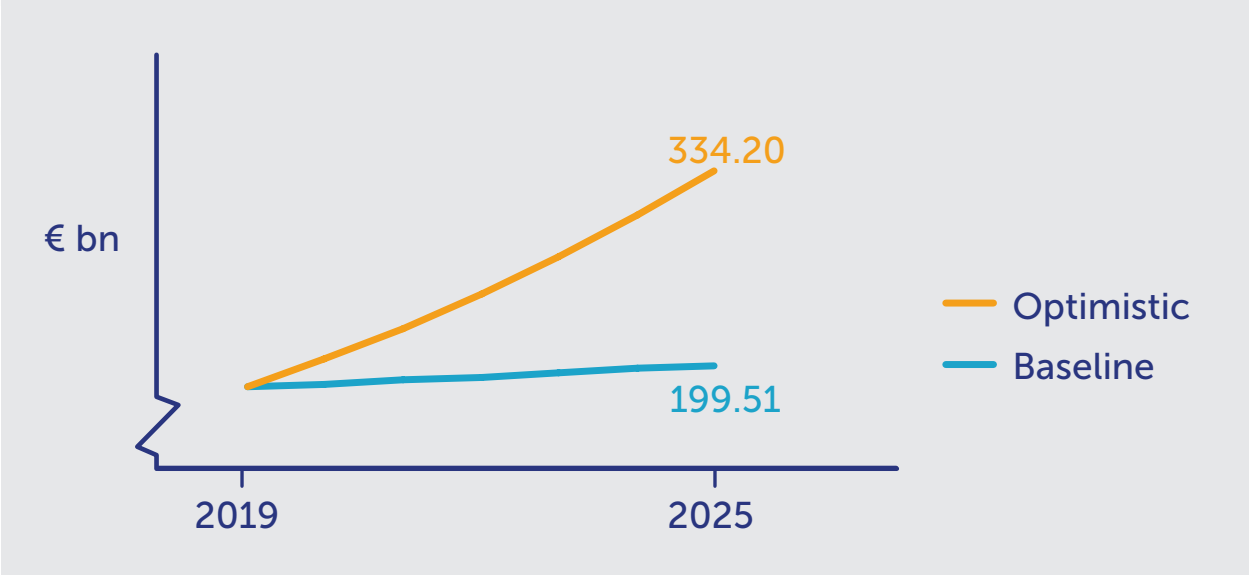
[14] European Data Portal (2015) available at: https://www.europeandataportal.eu/sites/default/files/edp_creating_value_through_open_data_0.pdf

Open data market size

2019

€184.45 billion market size

Open data market in 2019



€199.51 - €334.20 billion market size

Forecast for 2025 depending on the baseline (1.0-1.4%) or optimistic (10.4%) growth rates

2025

Current and forecast numbers are for EU27+ countries.
For details on calculations and assumptions see corresponding sections.

EUROPEAN DATA PORTAL

2. THE OPEN DATA MARKET SIZE

2.1 Method for measuring the open data market size

Measuring the market size is usually a rather straight forward task. The price of a product or a service can be multiplied by the amount of its sales. A potential market size can be assessed by estimating the demand for a certain product or service and its price.

Open data is available free of charge and its exact reuse is barely documented. In addition, reuse is endless, and the network effects make open data more impactful the more it is reused. Without being reused the data has no value and cannot create impact. **The market size of open data therefore also could be called the market size of products, services, and content improved or enabled by open data.** Which share of this value is attributable to open data can only be estimated. Therefore, measuring the market size of open data is more complex and less precise. There are multiple methodological approaches anyway. This is because it helps to compare the economic impact of open data to other contributors. It gives the abstract concept of open data as an economic factor a tangible value. This helps to formulate open data strategies to further exploit the value for Europe of its public data.

In the report, the market size is expressed in Euro and is derived from the Gross Domestic Product (GDP). This approach is similar to other studies measuring the open data market size, such as Lateral Economics & ODI (2016)[15] or Vickery (2011)[16]. GDP refers to the total market value of all final goods and services produced in an economy in a given year (sum of the Gross Values Added (GVA) of all resident producers at market prices, plus taxes minus subsidies on imports)[17].

To express the economic impact of open data as EU market size, we use secondary and primary data, and adapt literature results to make informed and meaningful statements about the current market size and the potential growth to forecast the open data market size for 2025. The steps undertaken are listed below:

1. Reviewing and reusing data from relevant literature.
2. Translation of results into share of GDP.
3. Equalizing the results from literature to make them comparable.
4. Calculation of the current open data Market size.
5. Forecast of the open data market size 2025 using different growth scenarios.

2.2 Relevant literature to measure the open data market size

For the literature research, we included research from 2000 onwards with a focus on more recent studies. We only include the most relevant literature dating from before 2013, as before then, the concept of open data was barely used - and still referred to as public sector information (PSI). However, many relevant studies after 2013 refer to or are based on approaches and or results from those early studies like PIRA (2000)[18] and MEPSIR (2006)[19] researching the value of PSI.

[15] Lateral Economics & ODI (2016) available at: https://theodi.org/article/research-the-economic-value-of-open-versus-paid-data/
[16] Vickery (2011) available at: http://ec.europa.eu/newsroom/document.cfm?doc_id=1093
[17] OECD (2002) https://stats.oecd.org/glossary/detail.asp?ID=1170
[18] PIRA (2000) available at: http://ec.europa.eu/newsroom/document.cfm?doc_id=1195
[19] MEPSIR (2006) available at: http://ec.europa.eu/newsroom/document.cfm?doc_id=1197

The literature includes macro and micro-economic approaches that aim to quantify the impact of open data (respectively PSI for studies before 2013) based on primary and secondary data. Macro-economic studies start from a high aggregation level like GDP and - top down - adapting those to the open data market. Details on the composition or origin of impact created can only be discussed but findings are by nature generalised and modelized. Micro-economic studies start at a medium or low level of aggregation like surveys or case studies and - bottom-up - extrapolate and pro-rate the impact of open data to EU market size (see DotEcon (2015)[20]; POPSIS (2011)[21]; McKinsey (2013)[22]). Origin and composition of impact is immediate but not mutually exclusive and collectively exhaustive (also known as the MECE principle).

The report includes studies with a supranational, national or sectoral focus. Studies focusing on niche markets or regions are excluded as they are hardly comparable or pro-rateable to the EU context. Selected studies differ in the scope of open data and in the scope of impact level (see chapter 1.3.2.). The report includes studies that focus on PSI, open government data, and open data. We acknowledge that those concepts are different, but as we can still benefit from the individual approaches and insights they are included in our research.

PSI vs open data and open government data

The main difference between the concepts is the commercial aspect. As by now, open data is characterised by being available free of charge or at marginal costs, the revenues for the public sector from selling PSI is no longer a reasonable measure to understand the value of open data (e.g. MEPSIR (2006)[23] and PIRA (2000)[24]).

Open data used to be treated almost congruent with open government data. However, this is changing as we look at data and data re-use more holistically and the concept of open data is also more and more applied in the private sector. The risk applying a wide scope to the concept of open data, like Deloitte (2013)[25] is that the impact level widens and a significant overlap with the impact of data and data sharing in general occurs. While a holistic view on data is highly beneficial, we are very clear about the context and the scope of data and impact we are referring to in each example or calculation.

Direct and indirect

In addition, we have to distinguish between direct market size and total market size. The direct market size refers to the monetised benefits that are realised in market transactions in the form of revenues and Gross Value Added (GVA). The indirect market size refers to the benefits of open data i.e. new job potential, new goods and services, time savings for users of applications using open data, knowledge economy growth, increased efficiency in public services and growth of related markets. This report will combine direct and indirect impact and refer to it as the total market size. Results from studies that focus only on the direct market size will be equalised to be comparable to those who focus on the total market size. A selection of studies with different methodological approaches that are most relevant for this report is introduced on the next page.

---

[20] DotEcon (2015) available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/418097/Evaluation_of_CUPI_study.pdf

[21] POPSIS (2011) available at: https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=1159

[22] McKinsey (2013) available at: https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/Open%20data%20Unlocking%20innovation%20and%20performance%20with%20liquid%20information/MGI_Open_data_FullReport_Oct2013.ashx

[23] MEPSIR (2006) available at: http://ec.europa.eu/newsroom/document.cfm?doc_id=1197

[24] PIRA (2000) available at: http://ec.europa.eu/newsroom/document.cfm?doc_id=1195

[25] Deloitte (2013) available at: https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/deloitte-analytics/open-data-driving-growth-ingenuity-and-innovation.pdf

The baseline to measure economic impact of PSI: PIRA (2000), MEPSIR (2006), and Vickery (2011)

PIRA (2000)[26] provided the first extensive research into the economic impact of PSI. The methodology in the PIRA study differs from more recent studies as the report used the investment value of PSI, that is government investment in the acquisition of PSI and the economic value expressed as part of national income from industries and activities that are based on the re-use of PSI. The latter is triangulated by expenditure of the private sector for PSI and the estimated value added by PSI re-users. The study focused on direct impact. In addition, it focuses on PSI like other studies before 2013 because concepts like open government data or open data were not coined yet.

MEPSIR (2006)[27] is the most comprehensive study of the European PSI market and the basis for many subsequent studies. The quality of the result is considered to be considerably higher than the more subjective estimates of open data market size. The report focuses on the direct market size of PSI in the European Union (EU25) and Norway and includes a comparison with the United States (US).

The methodologic approach is based on surveys of PSI suppliers and re-users. The median of €26.1 billion for 2006 was used and represents the sum of the turnover of all individual re-users, minus costs of acquiring public sector information from public content holders. This was equivalent to 0.25% of European GDP in 2006.

Vickery (2011)[28] connected existing literature and derived EU level results on the market size of open data. The study reviewed the existing literature about the open data market size and in addition used national and sectoral studies to extrapolate the results to EU market size.

Open data Market size realised by companies and their employees: ASEDIE (2019)

ASEDIE (2019)[29] focus on economic aspects of open data characterising of the Spanish infomediary sector (sector of organisations that re-use public data). The annual study is based on a survey of more than 500 infomediary companies, annual reports of the participating companies and the ASEDIE database. The research approach includes five aspects:

- An approximation of the infomediary activities and the market size
- A study on the business model and number of clients
- An analysis of the geographical area of commercial activities
- An overview of the type of clients
- A study on the evolution & opportunities in the infomediary sector

Open data market size realised by countries: e.g. Open Data Barometer 2018

Open Data Barometer (2018)[30] by the World Wide Web Foundation researched the effect of open data on the economy measured by economic growth. The sample selection of this study existed of 115 countries. The data was gathered via a government self-assessment, peer-reviewed expert survey with country specialists, dataset assessments, and 15 kinds of data per country.

[26] PIRA (2000) available at: http://ec.europa.eu/newsroom/document.cfm?doc_id=1195
[27] MEPSIR (2006) available at: http://ec.europa.eu/newsroom/document.cfm?doc_id=1197
[28] Vickery (2011) available at: http://ec.europa.eu/newsroom/document.cfm?doc_id=1093
[29] ASEDIE (2019) available at: http://www.asedie.es/assets/asedie.-infomediary-sector-report-2019.pdf
[30] Open Data Barometer (2018) available at: https://opendatabarometer.org/doc/leadersEdition/ODB-leadersEdition-Report.pdf

Open data market size realised by extrapolating causal chains: McKinsey

McKinsey (2013)[31] used the bottom-up approach using causal value chains of how open data can create economic impact. The results are pro-rated to European level. The results are comparably higher because often the potential economic impact and not the actual impact is explored. In many examples, the economic impact of a future best-case scenario of opening or sharing specific information is explored. This explains the very optimistic result of the overall market size of this study compared to all other studies in the field.

The European Data Market Study by the International Data Corporation (IDC) & Lisbon Council and DemosEuropa & WISE

The European Data Market Study by IDC (2017)[32] determined three potential scenarios displaying the growth rate of the European data market and economy. Several macro-economic and framework conditions were studied and the trends in these areas were examined. The desk research included forecasts of EU GDP growth and ICT spending growth, economic growth conditions, policy and regulatory conditions, data market dynamics, and global trends affecting all technology markets. The study researched the value of the data market, the number of data user enterprises, the amount of data companies and subsequently their revenues.

The European Data Market Monitoring Tool: Key Facts & Figures, First Policy Conclusions, Data Landscape and Quantified Stories by IDC (2019)[33] is the update of the European Data Market Study and other relevant studies. It shares the newest results around the 6 key indicators of the European data market monitoring tool: number of data professionals, number of data companies, revenues from data companies, data market size (direct), data economy (direct and indirect), data skills gap. The work from the IDC together with the Lisbon Council is a highly relevant reference point for this study. Although the results express the value of the general data market end economy and not specifically open data.

DemosEuropa & Warsaw Institute for Economic Studies (2014)[34] used an approach to understand the macro-economic impact of big and open data on the European economy. The paper provides policymakers with recommendations of how to maximise the economic opportunity. It expresses the data market size and as part of it the open data market size.

Value added by Open data and paid data: ODI and Lateral Economics (2016)

Lateral Economics (2016)[35] focuses on value-added associated with open data. Their report states that the value-added associated with open data varies between 0.4 and 1.4 percent of GDP. The ODI (2016)[36] also expressed a ratio between the potential growth of paid and open data with open data growing 0.05% of GDP more than paid data.

---

31 McKinsey (2013) available at: https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/Open%20data%20Unlocking%20innovation%20and%20performance%20with%20liquid%20information/MGI_Open_data_FullReport_Oct2013.ashx

32 IDC (2017) available at: https://a2528ba5-a-c3c32646-s-sites.googlegroups.com/a/open-evidence.com/download/repository/SMART20130063_Final%20Report_030417_2.pdf?attachauth=ANoY7crL55TcXaDjVhYlzdD2jBNz6-UaN35LRS8sfZ9bqquqBazhj8Xk5FFQaus6z1xdbaviMZQ0Wbjkhst09Wpv7a13pztoDD9FrCANHECqEHeODrZgkvAvEes5w0ZiwC5t8L0Ovpk_7CHJVVPmUZWeUiVCYQoRGYk_b5ln0ZeTP3TH_A2sJhNeFtAg8X0ap3CqXfqHFmJGtgZkkEr0bcyM0pbRtH7UZ76oxlk89xXYB3FhdJ096eQqfhBG6Cx0TBGI9VCSKQRE&attredirects=0; http://datalandscape.eu/webinars/how-power-data-can-drive-european-union-economy-european-data-market-study-and-monitoring

33 IDC (2019) available at: http://datalandscape.eu/sites/default/files/report/D2.6_EDM_Second_Interim_Report_28.06.2019.pdf

34 DemosEuropa and Warsaw Institute for Economic Studies (2014) available at: https://tech.eu/features/381/open-big-data-in-europe/

35 Lateral Economics (2016) available at: https://www.scribd.com/doc/309810679/Permission-granted-The-economic-value-of-data-assets-under-alternative-policy-regimes

36 IDC (2016) available at: https://theodi.org/article/research-the-economic-value-of-open-versus-paid-data/

2.3 The total open data market size in 2019

To be able to compare the studies, the individual results regarding the market size are expressed as percentage of GDP of the year the study concerns.

Table 3: Overview of relevant results from literature about the open data market size

Study	Result expressed in the study	Results related to year	Adapted result as % of GDP[37]
PIRA 2000[38]	€68 billion direct market size of PSI in EU15	1999	2.69%*
MEPSIR 2006[39]	€27 billion direct market size of PSI in EU27 and Norway	2005	0.91%*
Pollock 2010[40]	£4.5 billion in UK PSI data market size (welfare gains)	2007	1.09%*
Vickery 2011[41]	€140 billion EU27 PSI market size	2009	1.10%
Buergi Schmelz 2013[42]	2.2 billion CHF in Switzerland open government data market size	2013	1.27%*
Deloitte 2013[43]	£6 billion in UK open government data market size	2011	1.42%*
McKinsey 2013[44]	$3 380.8 billion global open data market size	(no year indicated)	4.26%
Lateral Economics 2014[45]	$22 billion (AUD) in Australia open data market size	2013	1.40%
DemosEuropa and WISE 2014[46]	€200 billion for data market and €10 billion open government data market in EU28	2020	0.07%

[37] Specific to year and geographical scope of each study
[38] PIRA (2000) available at: http://ec.europa.eu/newsroom/document.cfm?doc_id=1195
[39] MEPSIR (2006) available at: http://ec.europa.eu/newsroom/document.cfm?doc_id=1197
[40] Pollock (2010) The Economics of Public Sector Information; Available at: https://ses.library.usyd.edu.au/bitstream/handle/2123/6568/PSI_vol1_chapter3.pdf?sequence=1n
[41] Vickery (2011) available at: http://ec.europa.eu/newsroom/document.cfm?doc_id=1093
[42] Buergi Schmelz (2013): Wirtschaftliche Auswirkungen von Open Government Data; Available at: https://www.egovernment.ch/index.php/download_file/force/347/3337/
[43] Deloitt (2013) Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/198905/bis-13-743-market-assessment-of-public-sector-information.pdf
[44] McKinsey (2013) available at: https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/Open%20data%20Unlocking%20innovation%20and%20performance%20with%20liquid%20information/MGI_Open_data_FullReport_Oct2013.ashx
[45] Lateral Economics (2014) available at: https://www.omidyar.com/sites/default/files/file_archive/insights/ON%20Report_061114_FNL.pdf
[46] DemosEuropa and Warsaw Institute for Economic Studies (2014) available at: https://tech.eu/features/381/open-big-data-in-europe/

DotEcon 2015[47]	£1.1 billion direct PSI market size in UK	2005	0.29%*
EDP 2015[48]	€193 - €209 billion total open data market size in EU28+	2016	1.35%[49]
ODI and Lateral Economics (2016)[50]	Value added associated with open data varies between 0.4 and 1.4 percent of GDP	(no year indicated)	1.20%
ASEDIE (2019)[51]	Value of turnover from sales of PSI-based goods and services of €1.7 billion in Spain	2016	0.53%*
IDC Data Landscape (2019)[52]	Direct EU data market size €50.4 billion, indirect EU data market size €335.5 billion	2017	2.18%[53]
World Wide Web foundation 2018 (Open Data Barometer)[54]	€477.3 billion for EU28	2020	0.13%

Some studies focus only on the direct impact (see e.g. PIRA (2000)[55]). Therefore, we use a ratio to pro-rate the results for the total market size including the indirect market size. In that way we increase comparability of the results. Literature suggests ratios of for example 3.78 (Shakespeare (2013)[56]) or 3.50 (Vickery (2011)[57]). The results from the latest EU data market study even indicate a ratio of around 6.7 of open data to "paid data". We use the more conservative 3.64 (the average of Shakespeare and Vickery) to equalise the different levels of impact as indicated in the table with an asterisk (*). At the same time, we acknowledge and support the position that indirect impact lately is increasing relatively to the direct impact. This is considered when forecasting the open data market size for 2025.

Now that we can compare the shares in GDP, we can derive an average value. The average share of GDP is 1.33% and the median is 1.19%. Significant is that although the approaches vary from bottom-up top down, the results are in a rather small range between as shown in the figure. Moreover, the results do not testify to a clear increase of the open data market size as share of GDP over the years. This finding will get more important when forecasting the open data market size growth scenarios.

47 DotEcon (2015) Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/418097/Evaluation_of_CUPI_study.pdf
48 European Data Portal (2015) available at: https://www.europeandataportal.eu/sites/default/files/edp_creating_value_through_open_data_0.pdf
49 Not published in EDP (2015), but calculated as average of the total open data market size (€193 - €209 billion) as share of GDP (1.29% - 1.4%) in 2016. Note: the 0.35% and 0.47% published in the 2015 report only refer to the direct open data market size as share of GDP in 2015 and 2020 respectively. European Data Portal (2015) available at: https://www.europeandataportal.eu/sites/default/files/edp_creating_value_through_open_data_0.pdf
50 Lateral Economics & ODI (2016) available at: https://theodi.org/article/research-the-economic-value-of-open-versus-paid-data/
51 ASEDIE (2019) available at: http://www.asedie.es/assets/asedie.-infomediary-sector-report-2019.pdf
52 IDC (2019) Data Landscape; Available at: http://datalandscape.eu/
53 To derive the open data market size from the whole data market, we use a ratio of 5% of open data contribution to the overall data market. This ratio is used in literature, e.g. by demos WISE (2014). Open data experts from The Lisbon Counsel, the University of Southampton, the EU Open Data Portal and the European Data Portal confirmed this assumption to be realistic for a generic, simplified model.
54 Open Data Barometer (2018) available at: https://opendatabarometer.org/doc/leadersEdition/ODB-leadersEdition-Report.pdf
55 PIRA (2000) available at: http://ec.europa.eu/newsroom/document.cfm?doc_id=1195
56 Shakespeare (2013) available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/198752/13-744-shakespeare-review-of-public-sector-information.pdf
57 Vickery (2011) available at: http://ec.europa.eu/newsroom/document.cfm?doc_id=1093

The open data market size in 2019

To derive the approximate current open data market size the median of 1.19% is applied to EU GDP. The latest available official numbers are for 2018 were GDP was €17 939.94 billion for the EU28+[58]. The indicative GDP for 2019 is around €18 137.28 billion for the EU28+ (the growth rate of 1.1% indicated by the European Central Bank (ECB) in Q4 2019 was applied[59]). This leads to a total open data market size in 2019 for the EU28+ of around **€215.29 billion**. For the EU27+, i.e., excluding the UK, it is approximately **€184.45 billion**.

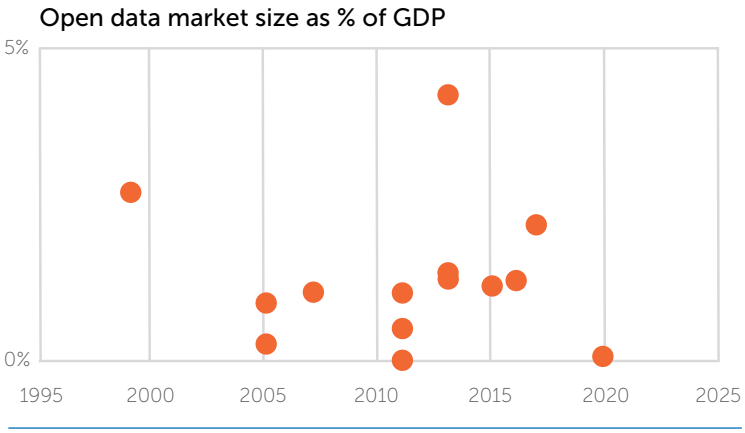


Figure 1: Visualisation of relevant results from literature about the open data market size

This is slightly lower than the total open data market size forecast for EU28+ by the 2015 EDP study on the economic impact of open data. There, the forecast was between €245 - €264 billion for 2019. This is mainly because the 1.35%[60] (not published in the EDP report (2015)) share of GDP used as a basis for the forecast in the 2015 report is higher than the one used in this report (1.19%). In 2015, the share was based on the MEPSIR study, while in this report we were able to include more literature results which suggest a lower GDP share. Contextualising the lower number indicates that the 2015 forecast was slightly too optimistic and efforts to increase the impact did not materialise as expected. Moreover, in the 2015 report, it was expected that 80% of the EU28+ would be in the highest open data maturity cluster (i.e., Trendsetter). However, in 2019 only three countries are in the Trendsetter cluster and eight in the second category of Fast-Trackers, in part due to revised measurement of maturity in the Open Data Maturity Report (2019)[61]. However, the aim is not to correct *ex post* or validate the 2015 forecast. Instead, the value lies in observing and contextualising different findings and their implications in order to express and assess strategies to increase the impact of open data.

Open data market size as part of the overall data economy

IDC (2019) estimates that the data economy (direct and indirect impact) will reach €388 billion in 2020 for the EU27. To compare our results about the total (direct and indirect) impact of open data, we use numbers for the EU27 (instead of EU27+ as provided above). For EU27, the total open data market size in 2019 is around €172.30 billion and is expected to reach €174.20 billion in 2020. Of those €174.20 billion in 2020, around €47.9 billion will be direct impact.

Comparing these numbers to the overall data economy suggests that the open data impact represents a significant part of the overall data economy. Around 12% of the value creation in the data economy is *directly* created by open data, the results suggest, and around 45% are *indirectly* impacted by open data.

[58] Eurostat available at: http://ec.europa.eu/eurostat/web/products-datasets/-/tec00118
[59] https://www.ecb.europa.eu/stats/ecb_surveys/survey_of_professional_forecasters/html/table_hist_rgdp.en.html
[60] Calculated as average of the total open data market size (€193 - €209 billion) as share of GDP (1.29% - 1.4%) in 2016. Note: the 0.35% and 0.47% published in the 2015 report refer only to the direct open data market size as share of GDP in 2015 and 2020 respectively. European Data Portal (2015) available at: https://www.europeandataportal.eu/sites/default/files/edp_creating_value_through_open_data_0.pdf
[61] EDP (2019) available at: https://www.europeandataportal.eu/sites/default/files/open_data_maturity_report_2019.pdf

These percentages seem very high. However, we have to consider that open data impact is not created stand-alone but is interlinked with the use of data from other sources and digital tools. Value creation of open data and of other data cannot simply be added. They are intertwined, which creates an overlap of impact created. In other words, 45% of value creation in the data economy is created partially by open data and 12% is created mainly by open data.

Artificially separating open data value creation from value creation due to other data is not purposeful nor realistic. Strategies to increase open data impact as well need to consider data in a holistic view including the bigger picture of digitisation and new (digital) ways of working and digital skills. Furthermore, the ODI also states that within the data economy the value creation due to open data increases faster than the general data economy[62] which ultimately leads to higher percentages of open data value creation within the data economy. This relation of open data as part of the data economy is further discussed in chapter 8.

2.4 The total open data market size in 2025

Based on the 2019 open data market size, a forecast for the growth until 2025 is calculated. To capture the spectrum of growth scenarios, we provide a baseline and an optimistic growth scenario.

2.4.1 Baseline scenario

The baseline scenario assumes that the impact of open data only grows at the same pace as GDP. The EU GDP growth rates range between -5.5% and 4.6% since 2000 and got more stable in the last five years between 1.4% and 2.9%[63]. In this scenario, no efforts towards increasing the impact of open data are taken into account.

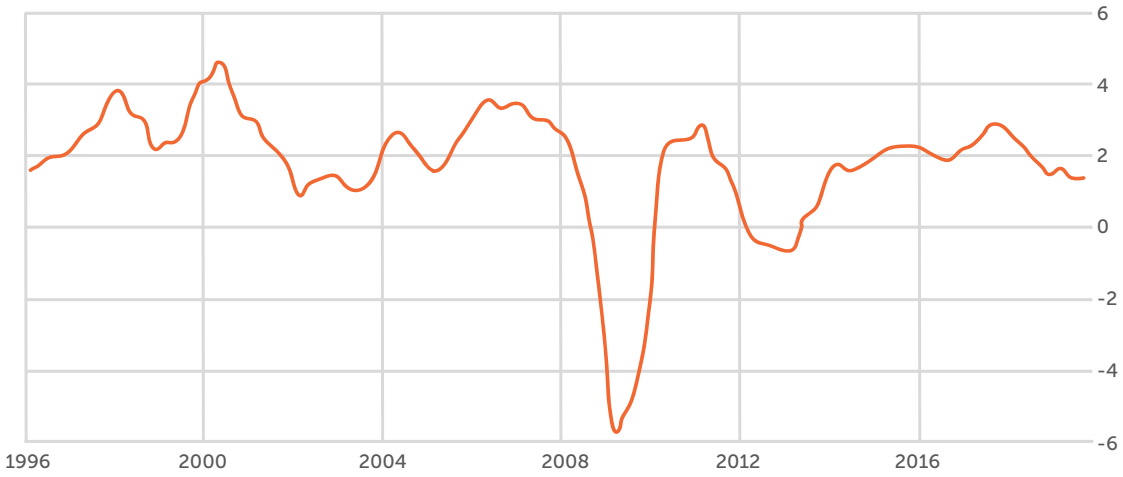


Figure 2: EU GDP growth; Trading Economics (2019);
available at https://tradingeconomics.com/european-union/gdp-annual-growth-rate

[62] ODI (2016) available at: https://theodi.org/article/research-the-economic-value-of-open-versus-paid-data/
[63] Trading Economics (2019) available at: https://tradingeconomics.com/european-union/gdp-annual-growth-rate

Real GDP growth rates (European Central Bank[64]) that include inflation corrections are used to forecast EU GDP, and at a stable level of 1.27% of EU GDP, the open data market size. For the subsequent calculations, we use the EU27+ open data market size from 2019 as a basis for our forecasts to take into account the withdrawal of the United Kingdom from the EU. As shown in the table, the open data market size in 2025 is forecast to be **€199.51** billion in a baseline scenario (€249.93 billion in EU28+).

Table 4: Baseline open data market size growth forecast

Year	2019	2020	2021	2022	2023	2024	2025
GDP in € billion for EU27+	15 539.24	15 694.63	15 898.66	16 121.24	16 346.94	16 575.80	16 807.86
ECB: expected real GDP growth in %	1.10	1.00	1.30	1.40	1.40	1.40	1.40
Baseline: open data market size in € billion for EU27+	184.45	186.30	188.72	191.36	194.04	196.75	199.51

2.4.2 Optimistic growth scenario

However, open data experts and several studies often use higher growth rates for the open data market size than GDP growth. Although calculations over the past years do not testify this growth potential for open data, there are several indicators that open data will indeed grow by a higher percentage than EU GDP. (These indicators are further discussed in the next chapters.)

- The new Open Data and PSI Directive and the specification and implementation of high-value datasets increase the economic impact of open data-driven services
- Increased uptake of data sharing which increases attention for open data in new target groups
- Habituation to GDPR and regain of confidence when reusing data
- Open data sectors and fields like language technology, agriculture and smart mobility drive open data growth
- Network effects increase and multiply the growth and increase specifically the indirect impact
- Increase in employment rates potentially leading to higher value creation.

Lateral Economics (2016)[65] states: "the values of open data (as a percentage of GDP) will increase over time due to the rapid expansion of new applications and the greater opportunities for re-use by consumers as a result of increased penetration of digital devices". The IDC's (2018)[66] EU Data Market Study projects a 0.05% higher growth in share of GDP in comparison to the overall data market that increased about 11.8% according to their study. Vickery (2011)[67] suggests a growth rate of around 7%, IDC (2019)[68] as part of the EU Data Market Study suggests three growth rate scenarios ranging between 4% and 16%.

64 European Central Bank (2019) available at: https://www.ecb.europa.eu/stats/ecb_surveys/survey_of_professional_forecasters/html/table_hist_hicp.en.html
65 Lateral Economics (2016) available at: https://www.scribd.com/doc/309810679/Permission-granted-The-economic-value-of-data-assets-under-alternative-policy-regimes
66 IDC (2018) available at: http://datalandscape.eu/european-data-market-monitoring-tool-2018
67 Vickery (2011) available at: http://ec.europa.eu/newsroom/document.cfm?doc_id=1093
68 IDC (2019) available at: http://datalandscape.eu/sites/default/files/report/D2.6_EDM_Second_Interim_Report_28.06.2019.pdf

In the table below the optimistic forecast for the open data market size is shown. Instead of using the forecast EU GDP growth rates, a growth rate of 10.413% (EU27+) is applied that takes into account data and open data specific aspects and trends. The growth rate is based on the growth scenarios from IDC (2019)[69] applied to the EU27+ and EU28+, based on their individual open data maturity and open data maturity trend. This results in an open data market size of around €334.20 billion in 2025. (EU28+ €253.74). The calculation is explained in more detail below.

Table 5: Optimistic open data market size growth forecast

Year	2019	2020	2021	2022	2023	2024	2025
Optimistic growth: open data market size in € billion, EU27+	184.45	203.66	224.86	248.28	274.13	302.68	334.20

2.4.2.1 Defining the growth rate

Previous literature often applied growth rates by Vickery (2011)[70]. These are based on three previous studies dating from the years before 2011. As the open data market has rapidly evolved, the prediction of 2011 might not be as reflective anymore. Therefore, the more recent growth rates from IDC are used to forecast the open data market size for 2025. The benefit of using the growth rates calculated by IDC (2019)[71] is that these figures are based on data gathered in the past years. We assume that several factors included in this study, such as data market dynamics factors, and global megatrends affecting all technology markets are significant for the open data market as part of the data market. The three growth scenarios are:

- A baseline scenario (7.5% growth rate): this scenario is based on the assumptions that the current growth trends and evolution would continue in a similar fashion. *"The Baseline scenario is characterised by a healthy growth of data innovation, a moderate concentration of power by dominant data owners with a data governance model protecting personal data rights, and an uneven but rather wide distribution of data innovation benefits in the society"* IDC (2019, page 10).

- A challenge scenario (4.3% growth rate): this scenario displays the situation where the data market grows more slowly than in the baseline scenario. This lower growth would be due to a less positive macro-economic context and less favourable framework conditions. *"The Challenge scenario (Digital Maze) is characterised by a low level of data innovation, a moderate level of data power concentration due to digital markets fragmentation, and an uneven distribution of data innovation benefits in the society"* IDC (2019, page 11).

- A high growth scenario (15.7% growth rate): this scenario states that the data market enters a faster growth trajectory, due to more favourable framework conditions. *"The High Growth scenario (data-driven reality) is characterised by a high level of data innovation, low data power concentration, an open and transparent data governance model with high data sharing, and a wide distribution of the benefits of data innovation in the society"* IDC (2019, page 11).[72]

[69] IDC (2019) available at: http://datalandscape.eu/sites/default/files/report/D2.6_EDM_Second_Interim_Report_28.06.2019.pdf
[70] Vickery (2011) available at: http://ec.europa.eu/newsroom/document.cfm?doc_id=1093
[71] IDC (2019) available at: http://datalandscape.eu/sites/default/files/report/D2.6_EDM_Second_Interim_Report_28.06.2019.pdf
[72] IDC (2019) available at: http://datalandscape.eu/sites/default/files/report/D2.6_EDM_Second_Interim_Report_28.06.2019.pdf

2.4.2.2 Individualising the growth rates for EU27+ based on open data maturity and development trends

Because not all countries have the same open data growth potential, we map the open data maturity per country and the maturity trend[73] with the three IDC growth scenarios. The trend is derived from the maturity scoring since 2005 and categorised as low growth countries, medium growth countries and accelerating countries. It must be said that even in low growth countries open data impact can be high, e.g. if the overall maturity is high but simply not growing fast. We use the following rules to allocate the most realistic growth rate per country:

Table 6: Allocation of growth rates based on countries open data maturity and trend

High-growth countries with a maturity above 50. Medium-growth countries with a maturity above 60.	High growth scenario	15.7%	Example: Czech Republic, Estonia or Spain
Medium-growth countries with maturity below 60. Low-growth countries with a maturity above 60.	Modest growth scenario	7.5%	Example: Austria, Norway or Latvia
Low-growth countries with a maturity below 60.	Low growth scenario	4.3%	Example: Hungary or Iceland

After mapping the countries to the most realistic growth rate we can calculate the average growth, which is 10.413% for the EU27+ (EU28+ 10.488%) resulting in an open data market size of around €334.20 billion in 2025. (EU28+ €253.74).

2.4.3 Growth potential for the open data market size until 2025

Combining the two growth scenarios, reveals a potential of €134.67 billion for EU27+ of open data market size growth. Implications and feasibility will be further explored in the following chapters.

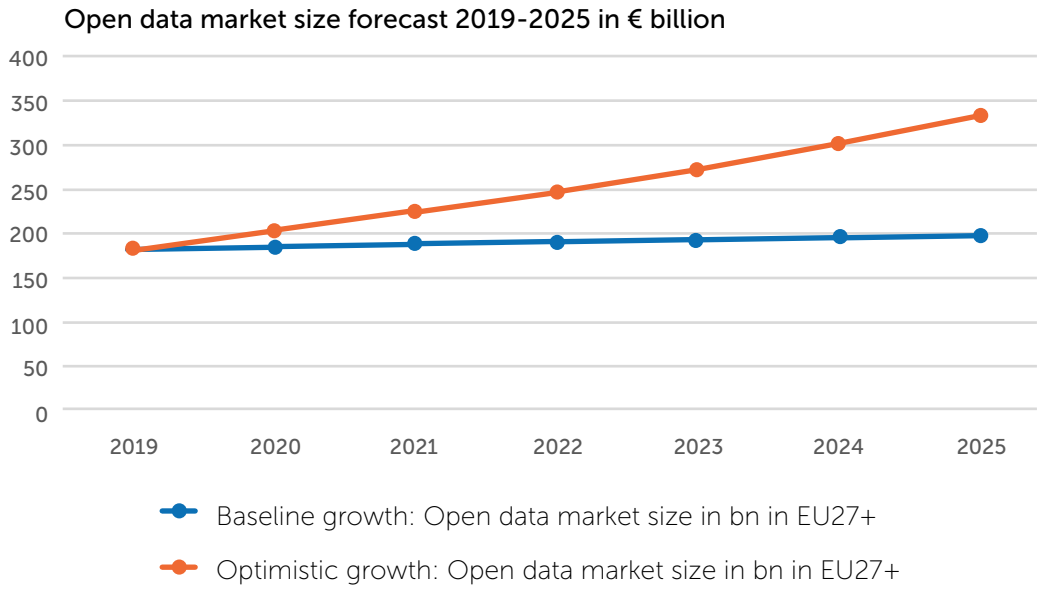


Figure 3: Visualisation of the baseline and the optimistic open data market size forecast

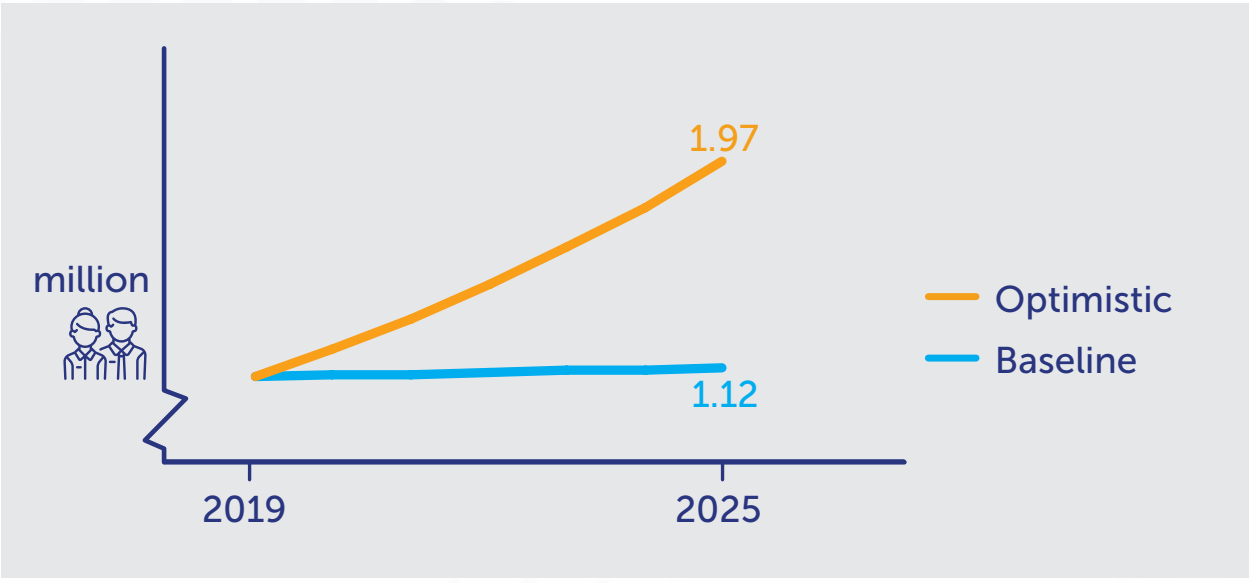
[73] EDP (2019) available at: https://www.europeandataportal.eu/sites/default/files/edp_landscaping_insight_report_n4_2018.pdf

Open data employment

2019

1.09 million employees

Open data employment in 2019



1.12 - 1.97 million employees

Forecast for 2025 based on the baseline (0.5%) or optimistic (10.4%) growth rates

2025

Current and forecast numbers are for EU27+ countries.
For details on calculations and assumptions see corresponding sections.

EUROPEAN DATA PORTAL

3. OPEN DATA EMPLOYMENT

Open data is now being recognised as an important asset for businesses and economic growth and as calculated in the previous chapter, the open data market size could in an optimistic scenario grow to **€334.20 billion** in 2025. The value of open data is created by both the public and the private sector. The economic impact of open data can - in addition to its market size - also be calculated in terms of the number of people who are employed due to open data. The potential for job creation through publishing and re-using open data is significant.

3.1 Relevant literature on open data employment

The number of studies available that measure the number of jobs created in the field of open data is relatively limited. Lateral Economics (2014)[74] argues that open data has the potential to transform labour markets through the transmission of information on workplace conditions to prospective employees. Open data on employment conditions would result in a better match of employees and organisations with the benefits of greater job satisfaction and improved productivity. Lateral Economics estimates that open data enabling better matching of people and jobs could add 3.4 billion AUD (~€2.1 billion) per annum to Australia's economy. When looking at the public sector, the report of the European Open Science Cloud expert group (2016)[75] stressed the need for hundreds of thousands of data experts to be trained by 2020 in order to realise open data in science. When looking at the private sector, the study by the European Data Portal into open data re-use (EDP (2015))[76] estimated that 25 000 jobs directly related to open data would be created in the private sector between 2016 and 2020. The Kyiv School of Economics researched the economic potential of open data for Ukraine. The study estimated that 3 thousand to 4 thousand employees worked directly with open data in Ukraine in 2018 and estimates a number of 4.2 thousand to 5.9 thousand open data employees in 2025[77]. In Spain, the multisectoral information association ASEDIE conducts annual research into Spanish infomediary companies. These companies re-use open data to create products and services. The latest study from 2019[78] found that these companies together employed more than 20 thousand people and generated €1.8 billion in 2017.

3.2 Method for measuring open data employment

Since open data is important for both the public and the private sector, we estimate the number of employees that are employed due to open data for both sectors. The number of employees is a more precise indicator than the number of open data jobs because jobs are measured in full-time equivalent (FTE) while employees are often not fulltime employed on open data jobs. This means that the number of jobs does not represent the actual number of persons that are employed due to open data, while the number of employees do. Therefore, we measure the number of employees (in person) and not the number of jobs that exist due to open data.

[74] Lateral Economics & Omydiar Network (2014) available at: https://www.omidyar.com/sites/default/files/file_archive/insights/ON%20Report_061114_FNL.pdf
[75] European Commission (2016) available at: https://ec.europa.eu/research/openscience/pdf/realising_the_european_open_science_cloud_2016.pdf
[76] European Data Portal (2015) available at: https://www.europeandataportal.eu/sites/default/files/edp_creating_value_through_open_data_0.pdf
[77] Kyiv School of Economics (2018) available at: http://tapas.org.ua/en/media/ekonomichnyj-potentsial-vidkrytykh-danykh-v-ukraini/
[78] ASEDIE (2019), available at: http://www.asedie.es/assets/asedie.-infomediary-sector-report-2019.pdf

There is a difference between employees that are directly employed due to open data and employees that are indirectly employed due to open data. Direct open data employees are people (not FTE) in both the public and the private sector, that are generating, providing, aggregating, re-using, and enriching[79] open data or substantially enable others to do so in organisations that base their business case or core focus on open data. These employees are directly employed due to open data because the start-up would not have existed without open data and therefore would not have employees. Another example is a civil servant who is employed to set up and maintain a national portal for the publication of open data. This employee is directly employed due to open data because its function would not have existed without open data.

Beyond directly employed persons, there are persons that are indirectly employed due to open data. These are people (not FTE) in both the public and the private sector, that are generating, providing, aggregating, re-using, and enriching open data in organisations that do not base their business case or core focus on open data. These employees would still be employed without open data, yet they create value with open data.

The objective is to forecast the future number of employees in the public and private sector that are employed due to open data. The first step is to estimate the number of open data employees in 2019. The second step is to forecast the number of open data employees in 2025. The private sector makes up 84% of the total employment in Europe[80].

3.3 Open data employment in 2019

The first step is to estimate the number of direct open data employees in 2019. The number of open data employees in Spain in 2017 serves as a starting point. The latest study from ASEDIE (2019)[81] provides insights into Spanish infomediary companies in 2017. These companies re-use public and/or private sector information to create products or services that add value. Since open data is at the core of the business models of these infomediary companies, we can use the latest study to estimate the number of employees that are directly employed due to open data. According to the study, 20 229 employees were employed in the 554 infomediary companies that were able to provide employee data in 2017. It is important to note here that only 74% of the companies in the sample provided employee data, which means that the actual number of employees is likely to be higher. For this report, we take a conservative approach and assume that there were just **20 229** direct open data employees in Spain in 2017.

3.3.1 Estimating direct open data employment in 2017

Now we know the number of direct open data employees in Spain in 2017, it is possible to calculate what percentage of the total employed persons in the country are directly employed due to open data. To do so, the total employment in Spain from Eurostat[82] is used. The table below shows the results of the calculation.

[79] Based on the Data Value Chain Archetypes of the European Data Portal (2015, p.29) available at: https://www.europeandataportal.eu/sites/default/files/edp_creating_value_through_open_data_0.pdf

[80] It should be noted that the limits of the public sector vary across Member States. For example, jobs in education or health are part of government employment in some countries, while they are not in others. Percentage based on Eurostat (2019) available at: https://ec.europa.eu/eurostat/cache/digpub/european_economy/index.html and https://ec.europa.eu/eurostat/web/products-eurostat-news/-/WDN-20170723-1

[81] ASEDIE (2019) available at: http://www.asedie.es/assets/asedie.-infomediary-sector-report-2019.pdf

[82] Eurostat definition employment: "all persons between 20 and 64 year who worked at least one hour for pay or profit during the reference week or were temporarily absent from such work".

Table 7: Share of open data employment in Spain

Country	Open data employment	Total employment	Share of open data employment
Spain	20 229	18 531 000	0.11%

In order to calculate the number of direct open data jobs in all other EU28+ countries, it is necessary to know the total number of employed persons in each of these countries and a reasonable share (%) than can be used. The total number of employed persons is derived from Eurostat[83]. The share of open data employees from Spain (0.11%), however, cannot be applied to all EU28+ countries because Spain is not an average country in terms of open data development: it had the second-highest open data maturity score of all EU28+ countries in 2017 according to the Open Data Maturity Report[84]. It is intuitive that there are more persons employed due to open data in a country with a higher open data maturity than in a country with lower open data maturity. Therefore, the share of 0.11% open data employment is a best-case scenario and will only be applied to countries that were - like Spain - considered to be open data trendsetters in 2017. Countries that fall into lower open data maturity groups than Spain, will likely employ less people due to open data. We estimate that countries in each of the following open data maturity groups - open data fast-trackers, followers, and beginners - employ 20% fewer employees due to open data than the superior group. The table below shows the share of open data employment (%) for each of the open data maturity groups.

Table 8: Share of open data employment per open data maturity cluster

Open data maturity cluster	Share of open data employment	Example
Trendsetters	0.11%	Ireland, Spain
Fast-trackers	0.09%	Cyprus, Poland
Followers	0.07%	Estonia, Lithuania
Beginners	0.06%	Liechtenstein

Based on the total employment numbers in each of the EU28+ countries and the above-mentioned shares, the number of direct open data employees in the countries is calculated. Adding the numbers of direct open data employees in each country together results in 225 thousand persons directly employed due to open data in EU28+ in 2017, and 192 thousand when we exclude the UK (EU27+) Knowing that the private sector represents 84% of total employment, we extrapolate the number to a total number of **229 thousand direct open data employees in EU27+ in 2019**.

3.3.2 Estimating total open data employment in 2017

Now we know the number of employees that are directly employed due to open data in 2017, we can calculate the number of employees that are indirectly employed due to open data. As an estimate, the earlier mentioned market size ratios of 3.78 by Shakespeare (2013)[85] or 3.50 by Vickery (2011)[86] can be

[83] Data derived from Eurostat, available at: https://ec.europa.eu/eurostat/web/products-datasets/-/t2020_10&lang=en
[84] European Data Portal (2017) available at: https://www.europeandataportal.eu/sites/default/files/edp_landscaping_insight_report_n3_2017.pdf
[85] Shakespeare (2013) available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/198752/13-744-shakespeare-review-of-public-sector-information.pdf
[86] Vickery (2011) available at: http://ec.europa.eu/newsroom/document.cfm?doc_id=1093

used. In this report, we use the average of both studies, i.e., a ratio of 3.64. By multiplying the number of directly employed persons with this ratio, it can be inferred that there are **833 thousand** persons indirectly employed due to open data. The number of direct and indirect open data employees together results in a total of **1.06 million persons employed due to open data in EU27+ in 2017**.

3.3.3 Translating open data employment from 2017 to 2019

In order to translate this number from 2017 to 2019, employment growth rates from the European Central Bank[87] and from Eurostat[88] are used, resulting in an employment growth of 1.4% for 2018 and 1.1% for 2019. Employment growth rates give the change in percentage from one year to another of the total number of employed persons. We assume here that the employment growth in open data in these years is similar to the employment growth in the EU28+ and EU27+. Applying these rates to the 1.06 million employees of the EU27+, results in **1.09 million open data employees in 2019**.

This is substantially higher than the open data employment forecast for EU28+ by the 2015 EDP study on the economic impact of open data (see page 81)[89]. There, the forecast was between 312 600 – 377 000 total open data employees for 2020. This shows that open data employment in 2019 is already higher than forecast for 2020 in EDP (2015). This is mainly because in the 2015 study, open data employment:

a. did not include the public sector
b. did not include support functions in open data-based organisations as direct open data employees.
c. did only partially include people who work with or for open data in organisations that cannot be called "open data organisations", i.e., that do not base their business case around open data or make it their core focus. These people, however, substantially contribute to the overall value creation by increasing productivity and efficiency and saving costs due to open data-based products, services or insights for their organisations.

Open data employment as part of the overall data employment

To provide context to this number, we can look at the total number of employees in the general data market in Spain. The European Data Market study defines general data workers as those who "*collect, store, manage and analyse data, as their primary activity*" (IDC (2019)[91]. The study estimates that there were 356 000 data workers in Spain in 2016. Looking at the 20 229 direct open data employees in the private sector in Spain results in a 5.7% employment share. This percentage is consistent with the earlier mentioned 5% contribution of open data to the overall data market in chapter 2.2.

To compare the open data employment to the overall data employment in the EU27 of 8.25 million according to IDC we use the EU27 open data employment of 1.06 million expected for 2020. This results in a 12.48% employment share of total open data employees as part of the European data employment. The employment share of direct open data employees as part of the European data employment is 2.69%.

---

87 European Central Bank (2019) available at: https://www.ecb.europa.eu/pub/projections/html/ecb.projections201912_eurosystemstaff~c7a91336cb.en.html#toc1

88 Data derived from Eurostat, available at: https://ec.europa.eu/eurostat/web/products-datasets/product?code=tesem040

89 European Data Portal (2015) available at: https://www.europeandataportal.eu/sites/default/files/edp_creating_value_through_open_data_0.pdf

90 There are 312 600 to 377 000 total open data employees forecasted for 2020 (see EDP (2015), page 81) and from figure 29 (see EDP (2015), page 80) it can be retrieved that approx. 22 000 to 27 700 additional jobs are created in 2020, so subtracting 22 000 to 27 700 from 312 600 to 377 000 gives 290 600 to 350 000 total open data employees in 2019. Note: the number of employees in figure 26 (page 79) of the 2015 report refer only to the direct open data employees. European Data Portal (2015) available at: https://www.europeandataportal.eu/sites/default/files/edp_creating_value_through_open_data_0.pdf

91 IDC & Open Evidence (2017, page 2010) available at: http://datalandscape.eu/study-reports/european-data-market-study-final-report

Again, like in chapter 2, we have to consider that open data impact is not created stand-alone, but people create value using open data interlinked with other data. In other words, 12.48% of data employees create value using open data and 2.69% of data employees create value almost exclusively with open data.

3.4 Open data employment in 2025

Starting point for forecasting open data employment in 2025 for the EU27+, is the number of open data employees in 2019 in the EU27+, that is the above mentioned 1.09 million. We forecast a baseline scenario and an optimistic scenario.

3.4.1 Baseline scenario

In the baseline scenario, we assume that the open data employment growth follows the overall employment growth in the EU27+. For this scenario, we use the employment growth rates provided by the European Central Bank[92], European Economic Forecast[93], and Cedefop[94] and apply this to the EU27+, which results in an average annual growth rate of 0.5%. The baseline scenario forecasts **1.12 million** open data employees in 2025 in the EU27+. This would mean that almost 33 thousand additional open data employees are needed between 2019 and 2025. The results of the calculation are shown in the table on the next page.

Table 9: Baseline open data employment growth forecast

Year	2019	2020	2021	2022	2023	2024	2025
Baseline: number of persons employed due to open data in EU27+ (in thousands)	1 089	1 094	1 100	1 105	1 111	1 116	1 122

3.4.2 Optimistic growth scenario

As explained in chapter 2 of this report, there are several indicators that open data can grow by a higher percentage than the GDP in Europe. For example, the new Open Data and PSI Directive and the specification and implementation of high-value datasets could boost the open data market size in the coming years. It is likely to assume that a bigger open data market size will also employ more employees. For the optimistic scenario, we assume that the open data employment grows with a higher percentage than the overall employment growth. This assumption is supported by the open data employment growth in Spain: the open data employment grew by 6.6% between 2011 and 2017[95], although overall employment in Spain decreased with -0.3%[96], in the same years. For the optimistic scenario, we use the same growth rate that we used to calculate the optimistic scenario of the open data market size (see chapter 2.4.2 in this report), which is an annual growth rate of 10.413%. This includes the assumption that the value created by employee stays stable. The alternative assumption will be discussed in chapter 3.6. The optimistic scenario forecasts **1.97 million open data employees in 2025 in the EU27+**. This would mean that **883 thousand additional open data employees** are needed between 2019 and 2025. The results of the calculation are shown in the table below.

92 European Central Bank (2019) available at: https://www.ecb.europa.eu/pub/projections/html/ecb.projections201912_eurosystemstaff~c7a91336cb.en.html#toc2
93 European Economic Forecast (2019) available at: https://ec.europa.eu/info/sites/info/files/economy-finance/ip115_en_0.pdf
94 Data derived from Cedefop, available at: https://www.cedefop.europa.eu/en/publications-and-resources/data-visualisations/skills-forecast
95 ASEDIE (2019) available at: http://www.asedie.es/assets/asedie.-infomediary-sector-report-2019.pdf
96 Data derived from Eurostat, available at: https://ec.europa.eu/eurostat/databrowser/view/tesem040/default/table?lang=en

Table 10: Optimistic open data employment growth forecast

Year	2019	2020	2021	2022	2023	2024	2025
Optimistic growth: number of persons employed due to open data in EU27+ (in thousands)	1 089	1 202	1 327	1 465	1 618	1 786	1 972

The figure below shows the forecast for the baseline as well as the optimistic scenario.

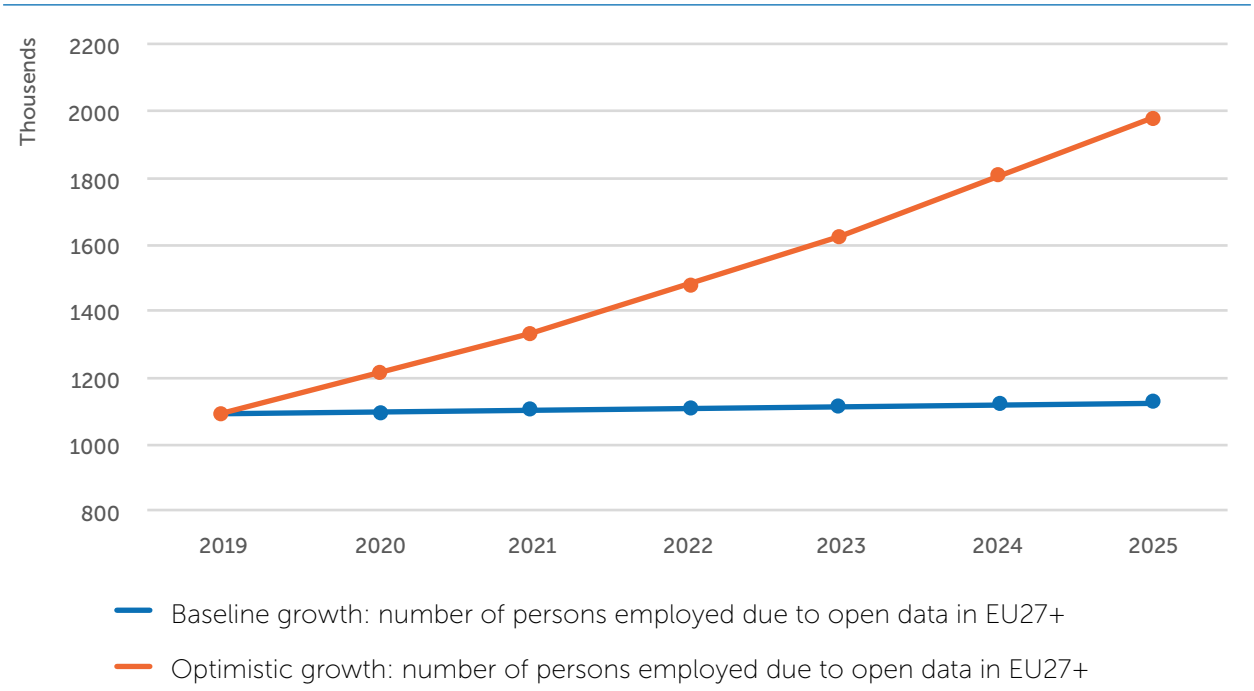


Figure 4: Visualisation of the baseline and the optimistic open data employment forecast

3.5 Open data employment in national governments

In order to gain benefits from open data, it first needs to be available and discoverable for everyone. In this chapter, we focus on the employment that is created by making open data available and discoverable on national level.

Open data can be made available and discoverable in many ways, such as on dedicated portals of public institutions or on regional, national, or European open data portals. National open data portals make open data from public and institutions and government bodies in the country discoverable. All the EU28+ countries except for Liechtenstein have a national open data portal. These national open data portals are set up and maintained by a national open data team. These teams are often not only responsible for the portal itself, but also for the coordination of open data initiatives in the country, the creation of supporting materials to publish and re-use open data, the provision of trainings and workshops to civil servants, and the promotion of open data publication and re-use through, for example, open data events.

The size of these **national open data teams** provides insight into the number of employees that are directly employed due to open data. As part of the open data maturity assessment of 2019[97], national open data teams provided information on the size of their team, more specifically on the number of persons working in their team and the number of full-time equivalents (FTE). In the EU28+, **100 employees** are directly employed due to open data. In the EU27+ **97 employees** are directly employed due to open data. It is important to note that seven countries did not provide information on the number of persons working in the national open data team. Therefore, the total number of employees in the EU27+ is likely to be higher. On average, **4 employees** are working per national open data team in the EU27+.

When looking at the number of FTEs, a total of **47 FTE** can be observed in the EU28+ and a total of **45 FTE** can be observed in the EU27+. The average number of FTE per national open data team is **2 FTE**. Comparing the average of 2 FTE with an average of 4 employees per national open data team, shows us that open data employees are often not full time employed. A potential explanation could be that employees do not exclusively work on open data but perform it as just one of their roles. Many governments work on data strategy and policy, which can include topics as data sharing, data management, and more general data topics. It is therefore likely that open data employees are involved in these topics in addition to their work related to open data.[98]

3.6 Value created by open data employees

It is estimated that there are 1.09 million open data employees in the EU27+ in 2019. Bringing these numbers into context with the different growth scenarios of the market size helps us to understand the open data employment and value creation better. The EU27+ open data market size has a value of €184.45 billion in 2019 (see chapter 2). Based on these numbers, we can estimate that on average, each open data employee creates a value of **€169 thousand** in 2019. Because we based the open data employment forecast on the assumption that the value creation per employee stays stabile, also in 2025 when the open data market size is predicted to reach a value of €334.20 billion in 2025, the forecast 1.97 million employees would still each create a value of €169 thousand in 2025. The question is if the relationship: "more employees, more total value created" is the most likely. Actually, value creation per employee is likely to change. An increase would lead to either a smaller workforce fulfilling the optimistic market size potential, or the forecast increase in workforce would create even more value.

97 European Data Portal (2019) available at: https://www.europeandataportal.eu/sites/default/files/edp_landscaping_insight_report_n4_2018.pdf
98 European Data Portal (2019) available at: https://www.europeandataportal.eu/sites/default/files/edp_landscaping_insight_report_n4_2018.pdf

Open data potential per sector

15.7% growth expected from specific sectors to reach optimistic market size

15.7%

High impact sectors

- Public administration
- Professional, scientific & technical
- Information & communication
- Transportation & storage

High potential sectors

- Agriculture
- Finance & insurance
- Health
- Education
- Retail & trade
- Real estate activities

The above estimations are potential numbers for EU28(+) or EU27(+) depending on availability of data. For details on calculations and assumptions see corresponding sections.

EUROPEAN DATA PORTAL

4. OPEN DATA POTENTIAL PER SECTOR

When forecasting the market size, we see a gap of around €135 billion between the baseline and the optimistic growth scenario. In order to reach this ambitious growth scenario, economic impact has to be created in different sectors. In the EU, the following sectors make up for over 80% of GDP[99]:

- Industry
- Wholesale/Transport/food service activities
- Public administration/education/health
- Professional, administrative and support activities
- Real estate activities

However, open data does not benefit each sector the same. The following review of relevant information supports a better understanding of open data impact in different sectors and domains.

4.1 Review of relevant input to explore the open data potential in different sectors

Sector assessment based on open data availability, downloads, and applicability: DemosEuropa & Warsaw Institute for Economic Studies (2014) and Deloitte (2013)

DemosEuropa & Warsaw Institute for Economic Studies (2014)[100] lists the most relevant sectors for the data economy:

- ICT with Finance and insurance
- Trade and manufacturing
- Public administration and health

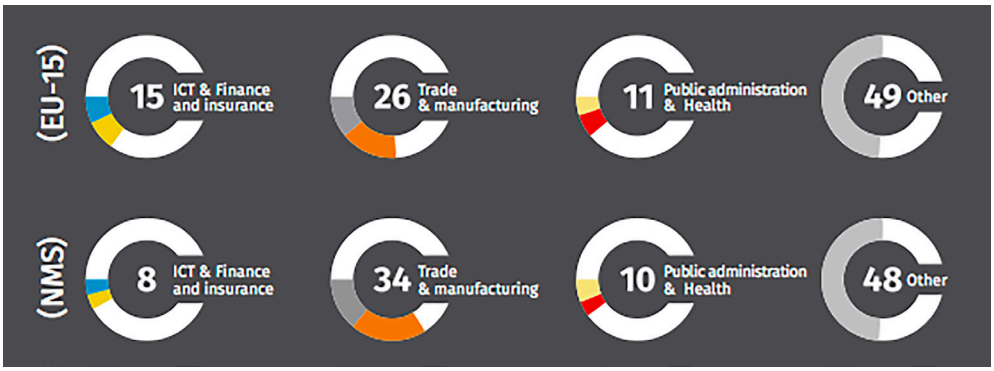


Figure 5: Most relevant sectors for the data economy; DemosEuropa & Warsaw Institute for Economic Studies (2014, page 53); available at: https://tech.eu/features/381/open-big-data-in-europe/

The same study also indicates the most important open data sectors:

- Public administration
- ICT
- Financial and insurance
- Real estate
- Health and social work
- Professional services

99 Eurostat (2015) available at: https://ec.europa.eu/eurostat/news/themes-in-the-spotlight/gva-employment

100 DemosEuropa and Warsaw Institute for Economic Studies (2014) available at: https://tech.eu/features/381/open-big-data-in-europe/

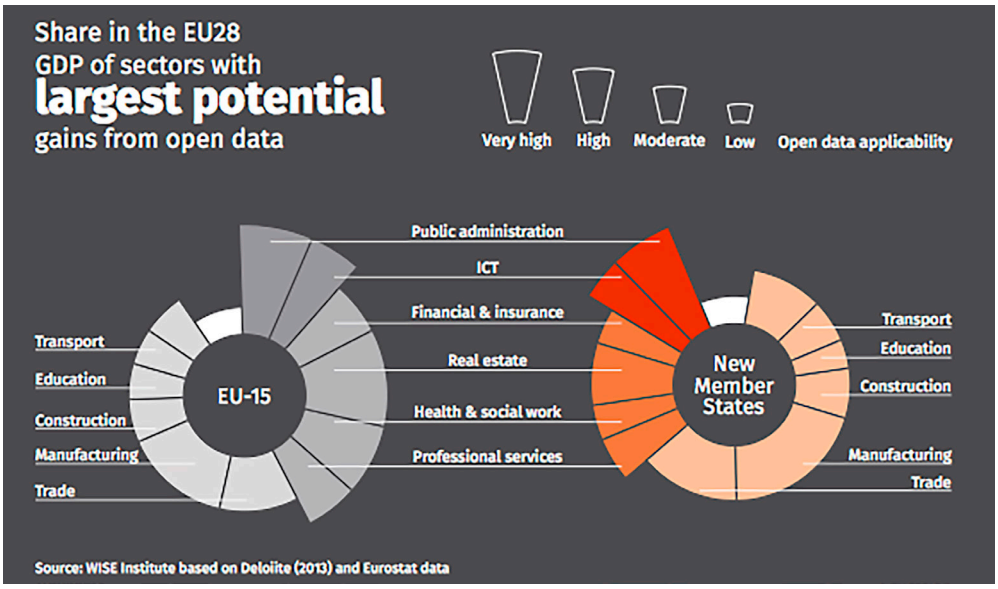


Figure 6: Sectors with the largest potential gains from open data; DemosEuropa & Warsaw Institute for Economic Studies (2014, page 72); available at: https://tech.eu/features/381/open-big-data-in-europe/

This indicates that for public administration and real estate, open data is especially beneficial compared to data in general. The study based this statement on results from Deloitte (2013)[101]. The study from Deloitte shows which sectors can potentially reuse most categories of open government data. This, however, has limited insight on the economic impact. Economic impact can be created by just one dataset while data from more categories does not necessarily lead to higher potential or higher impact. In addition, the results do not include actual reuse. Data that could benefit a sector but is not reused, e.g. because the quality is not sufficient, does not create economic value in the sector yet but has high potential.

Open data demand in different sectors: Open Data Barometer and EDP

Looking at the current demand side, the Open Data Barometer 2018[102] found out that map data, public transport timetables, and data on international trade and crime are among those datasets most demanded in the private sector. For the public sector, public data on health, education and environmental management is most demanded. This correlates with the findings from the report on PSI re-use in the public sector (2018)[103] where open data experts across Europe see the highest demand in maps and geodata, transport data, weather data, company registers, government procurement and spending, financial data, data on air quality, water quality, pollution, cadastre data and purchasing data. According to the report, the following domains will be most influenced by open data.

- Real-time data on transport and urban infrastructure in mobility, e.g. for smart transport, city maintenance and mobility as a service.
- Financial data for public administration, e.g. to support detecting fraud.

[101] Deloitte (2013) available at: https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/deloitte-analytics/open-data-driving-growth-ingenuity-and-innovation.pdf
[102] Open Data Barometer (2018) available at: https://opendatabarometer.org/
[103] European Data Portal (2018) available at: https://www.europeandataportal.eu/sites/default/files/analytical_report_11_psi_re-use_in_the_public_sector.pdf

- Public procurement data in public administration, e.g. for improved decisions making and spending, monitoring and assessing decisions and guide policymaking.
- Statistical sociocultural data, combined with anonymised personal health data, e.g. for analysing the relation between sleep patterns, weight, behaviour or fitness.

Open data re-use examples from different sectors: EDP and McKinsey

On the European Data Portal, a library of over 500 case studies of open data re-use are assembled. Case studies from the domains transport and government and public sector are the clear front runners in terms of number of case studies available.

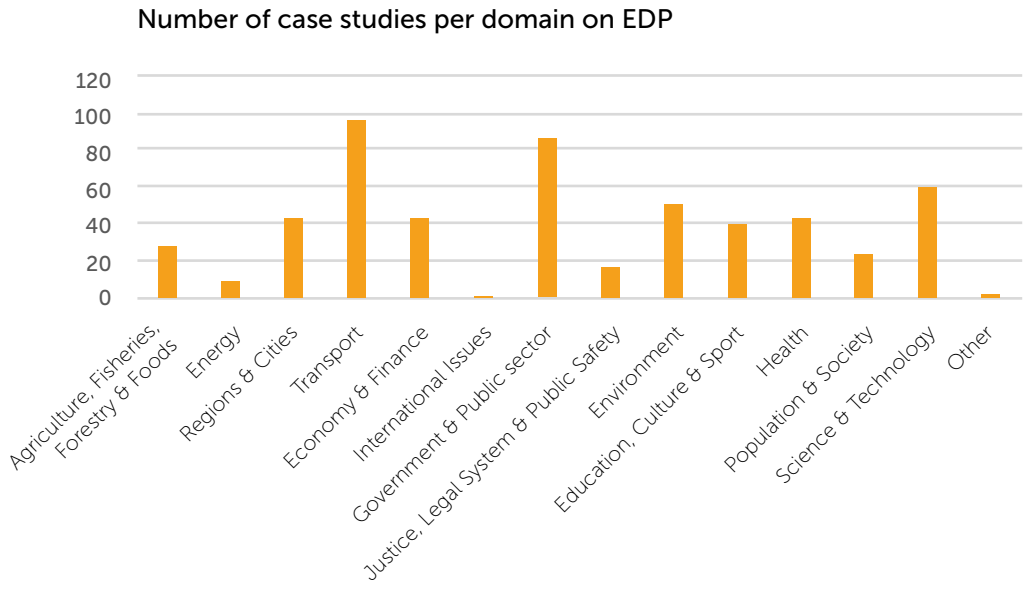


Figure 7: Overview of number of case studies available at the European Data Portal

The study "Open data: Unlocking inno-vation and performance with liquid information" from McKinsey (2013)[104] explored case studies and causal chains in different domains and also derived a shortlist of highly impacted sectors by open data. Those domains alone can create an economic impact of open data, according to the study, of €2.9 - €4.9 trillion ($3.2 - $5.4 trillion)[105]:

- Education
- Transportation
- Consumer products
- Electricity
- Consumer products
- Electricity
- Oil and gas
- Healthcare
- Consumer finance

[104] McKinsey (2013) available at: https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/Open%20data%20Unlocking%20innovation%20and%20performance%20with%20liquid%20information/MGI_Open_data_FullReport_Oct2013.ashx

[105] Exchange rates as of 26 January 2020

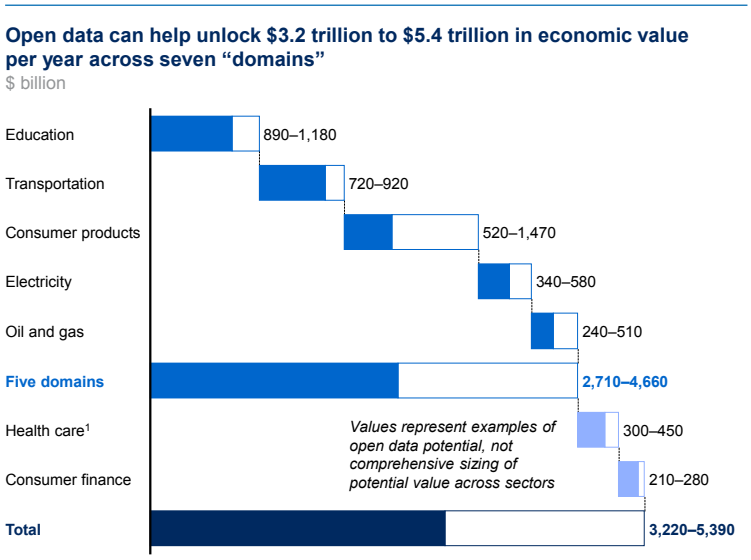


Figure 8: Domains in which open data can help unlock economic value; McKinsey (2013, page 9); available at: https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/Open%20data%20Unlocking%20innovation%20and%20performance%20with%20liquid%20information/MGI_Open_data_FullReport_Oct2013.ashx

Actual re-use number: results of the original field research for this report

In our survey (see chapter 7 of this report) participants indicated that more than half of the re-used data is about government and public sector (58). This could be due to the fact that people link open data to open government data and therefore indicate they re-used mostly data about government and public sector. However, to test this, we asked participants before about the concept of open data and this link was not prevalent so we can assume the re-users actually re-use open data about government and public sector and not from. The list is followed by regions and cities (40), population and society (37), science and technology (32), and transport (32).

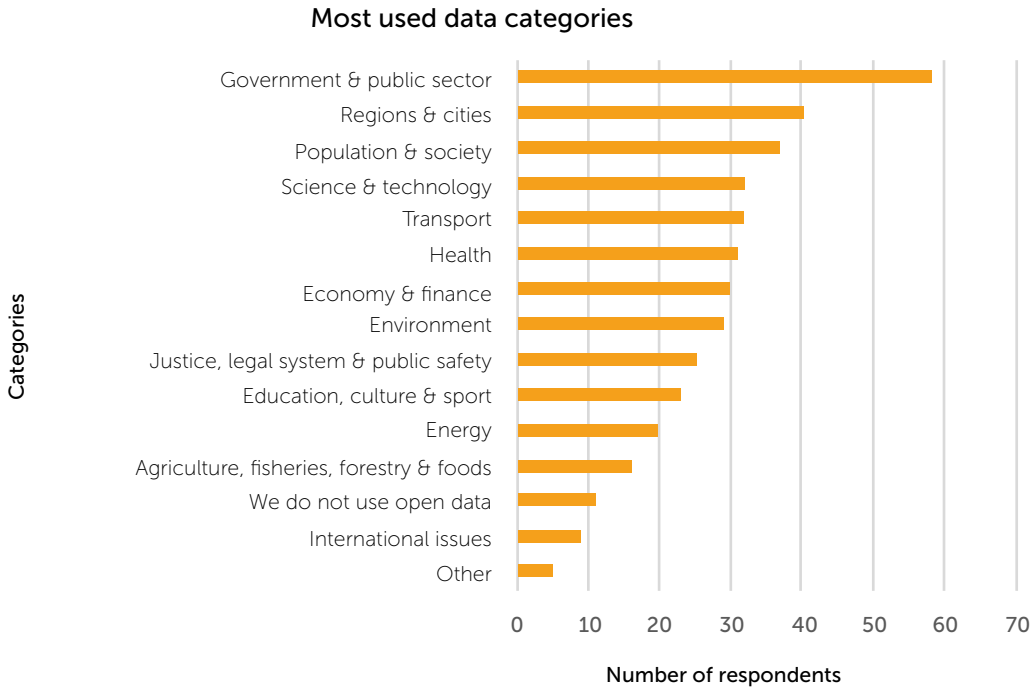


Figure 9: Most used data categories by respondents (in absolute numbers)

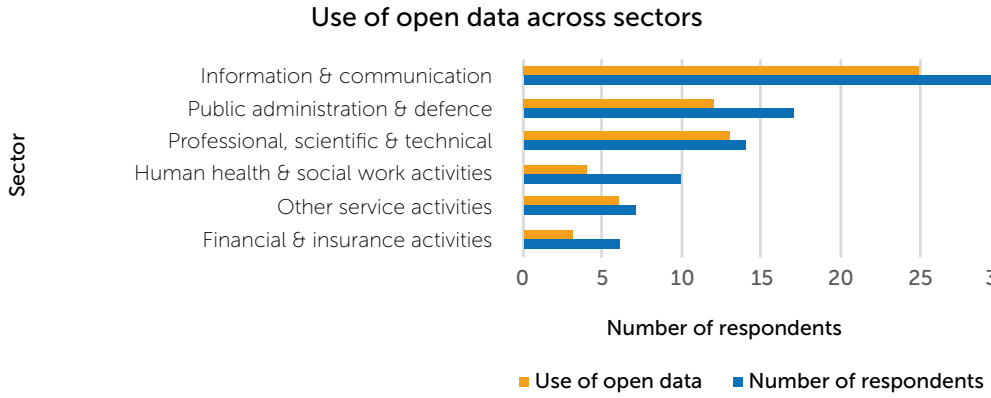


Figure 10: Use of open data across sectors (in absolute numbers)

4.2 Assessing open data potential per sector

Based on the findings from research, insights into digitisation of sectors in general (see e.g. Price Waterhouse Coopers (2011)[106]; Harvard Business Review (2016)[107]; McKinsey (2015)[108]; Digital Marketing Institute (2018)[109]) and experts insight into the matter, we categorise the EU sectors (NACE) in three groups based on their contribution and growth potential for the open data market size. For the three groups we, again use our growth rates individualised per country and its open data maturity and trend of 10.413% for EU27+:

- Group 1 - Growth scenario: very high 15.7%
- Group 2 - Growth scenario: high 7.5%
- Group 3 - Growth scenario: moderate 4.3%

The selection is based on three indicators:

- A - Level of digitisation and data demand
- B - Opportunity of open data supply to meet the demand
- C - Economic impact of the potential

[106] Harvard Business Review (2016) available at: https://hbr.org/2016/04/a-chart-that-shows-which-industries-are-the-most-digital-and-why

[107] McKinsey (2015) available at: https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/digital-america-a-tale-of-the-haves-and-have-mores

[108] Digital Marketing Institute (2018) available at: https://www.8pillars.com.au/wp-content/uploads/2018/09/eBook-Which-are-the-most-digital-industries-and-why.pdf

[109] Taken from the growth–share matrix, created by Henderson (1970) available at: https://www.bcg.com/publications/2014/growth-share-matrix-bcg-classics-revisited.aspx

Table 11: Group 1 - Contribution and growth scenario: very high 15.7%

Open data potential in	A	B	C	Sum
Agriculture, forestry and fishing	2	3	3	8
Wholesale and retail trade; repair of motor vehicles and motorcycles	3	2	3	8
Transportation and storage	3	3	3	9
Information and communication	3	2	3	8
Financial and insurance activities	3	2	3	8
Real estate activities	2	3	3	8
Professional, scientific and technical activities	3	2	3	8
Public administration and defence; compulsory social security	2	3	3	8
Education	3	2	3	8
Human health and social work activities	3	2	3	8

Table 12: Group 2 - Contribution and growth scenario: high 7.5%

Open data potential in	A	B	C	Sum
Manufacturing	2	1	2	5
Electricity, gas, steam and air conditioning supply	2	2	3	7
Water supply; sewerage, waste management and remediation activities	2	2	3	7
Accommodation and food service activities	2	2	3	7
Administrative and support service activities	3	2	2	7
Arts, entertainment and recreation	2	1	1	4
Other service activities	2	2	2	6
Activities of extraterritorial organisations and bodies	3	0	2	5

Table 13: Group 3 - Contribution and growth scenario: moderate 4.3%

Open data potential in	A	B	C	Sum
Mining and quarrying	1	1	2	4
Construction	1	2	1	4
Activities of households as employers; activities of households for own use	1	1	1	3

The average growth in all sectors based on the different growth scenarios is 10.9%. If the sectors indeed grow as expected, the optimistic open data market growth is feasible. That means, especially for the group 1 sectors, that the open data potential has to be leveraged and maximised.

The group 1 sectors (contribution and growth scenarios: very high 15.7%) can be split into two sub-groups, "high impact sectors" with proven and successful open data impact creation and the "High potential", that have a high potential to leverage more open data and thus create more value[110]. Especially in the sub-group of the stars, the opportunity for open data value creation needs to be strengthened to enable the growth potential.

High impact sector:
Public administration; Professional, scientific and technical activities; Information and communication and ICT; Transportation and storage
High potential sectors:
Agriculture; Financial services and insurance; Health; Education; Wholesale retail and trade; Real estate activities

It is also important to say that open data will play a significant role in the environment sector. Vickery (2011)[111], based on Craglia et al., already states that the economic impact of open data in the environmental sector was around €1 billion. With the recent increased interest and awareness for environmental topics, this sector - though not a separate NACE sector - is worth mentioning as a clear "star" open data sector. Value created for the environmental domain is further explored in the next chapter of this report (5.4). There causal value chains in selected promising domains are exemplified.

4.3 Value creation per employee differs per sector

Knowing which sectors are key in open data market size growth, enables us to specify opportunities and implications for growth. An increase in workforce can lead to more value created. A higher expected growth can make a bigger workforce needed. However, this causality is of course not generally applicable and especially for some sectors this causality is less true than for others. In the previous chapter, it was already indicated that value creation per employee is unlikely to stay stable. In order to provide context to the value that each open data employee creates, it is useful to look at the average revenues per employee in different kinds of companies and sectors. It is also observable that the average value creation per employee differs per sector, depending on the labour intensity. Low labour-intensive sectors, such as the ICT sector, tend to have higher revenues per employee due to higher amounts of automation and the use of new technologies. High labour-intensive sectors, such as agriculture, tend to have lower revenue per employee due to the physical efforts that are needed to complete necessary tasks[112].

When looking at the revenues per employee of eight leading tech companies, averages range from $360 thousand (~€326.5 thousand[113]) per employee up to $2.2 million (~€2 million[114])[115]. These tech companies

[110] Taken from the growth–share matrix, created by Henderson (1970) available at: https://www.bcg.com/publications/2014/growth-share-matrix-bcg-classics-revisited.aspx
[111] Vickery (2011) available at: http://ec.europa.eu/newsroom/document.cfm?doc_id=1093
[112] Investopedia (2019) available at: https://www.investopedia.com/terms/l/laborintensive.asp
[113] Exchange rates as of 26 January 2020
[114] Exchange rates as of 26 January 2020
[115] Statista (2019) available at: https://www.statista.com/statistics/217489/revenue-per-employee-of-selected-tech-companies/

have relatively high revenues and low labour-intensive jobs, which results - amongst other causes – in high revenues per employee. Looking at software companies worldwide, statistics show that the average revenue per employee is $173 thousand (~€157 thousand[116])[117], which comes close to the estimated €169 thousand that open data employees expected to have created in 2019.

Statistics from Eurostat (see figure on next page) support this statement, comparing Gross Value Added (GVA) per sector and employment in general (not open data specific), as percentages of the total. Financial and insurance activities show a similar trend of high GVA and comparably low numbers of employees. This results again in high-value creation per employee. Similar figures can be observed for the real estate sector and the information and communication sector.

When forecasting the sector growth and linking it to the open data employment, it is important to conclude that three out of the ten high growth sectors show higher value creation per employee and therefore offer less employment potential than the other seven sectors, where the GVA is realised by significantly more employees.

ASEDIE (2019)[118] states that the financial sector is among the fastest-growing sectors in terms of open data value creation, using indicators like new companies founded, sales generated and turnover growth. It is interesting to see in the study results, that the turnover per employee is 30% higher than the average. In the culture sector for example, direct value creation per employee is around €34 thousand (indirect value creation per employee expected around €123.76 thousand), less than 40% of the average of €88.8 thousand (indirect value creation per employee expected around €323 232 thousand). This underlines that also in open data related jobs and activities value creation per employee differs per sector.

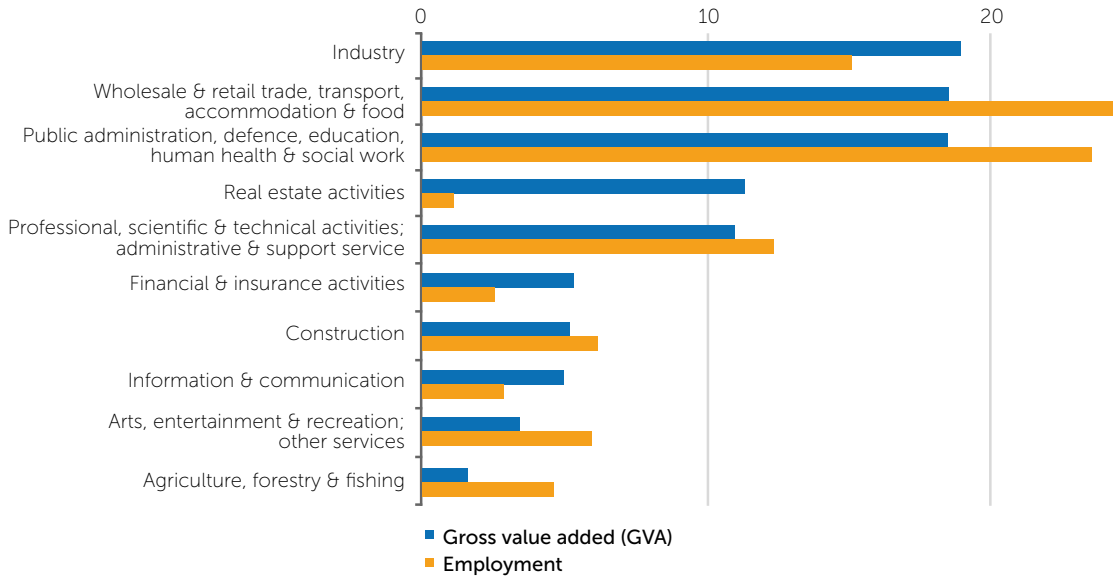


Figure 11: Gross Value Added (GVA) and employment per sector; Eurostat (2015)[119]; available at: https://ec.europa.eu/eurostat/news/themes-in-the-spotlight/gva-employment

[116] Exchange rates as of 26 January 2020
[117] Statista (2019) available at: https://www.statista.com/statistics/936887/software-professional-services-worldwide-revenue-per-employee/
[118] ASEDIE (2019) available at: http://www.asedie.es/assets/asedie.-infomediary-sector-report-2019.pdf
[119] Data displayed in the visualisation are the latest data available in the Eurostat online database. Figures may not add up to 100% due to rounding. *Activities of membership organisations, repair of computers, personal/household goods, other personal service activities.

5. EFFICIENCY GAINS DUE TO OPEN DATA

5.1 Intro and methodology

The previous sections have focused on the direct, monetised benefits of re-using open datasets in the form of revenues and gross value added, the increase in jobs generated, and cost savings. This section will elaborate on the indirect economic benefits, i.e., the efficiency gains, that come from new goods and services that in some way or form are beneficial to society. Efficiency in itself is a paramount factor in administrations, companies, and society in general. 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. In other words, we want to have the same things done faster or at a lower cost. Or in parallel, at the same time and cost we want to have a better outcome. The role of open data in enabling efficiency gains through re-use is the focus of this chapter.

Openly sharing data enables faster and easier access to information. As decisions are based on available information gathered from many different sources, one of the main benefits of open data sharing is better decision making, also known as "data-driven decision making". For example, data is used by individuals to decide which car to buy, which holiday to book, or which education to choose. The level of individuals' reliance on data, i.e., how they are taking advantage of data-driven solutions in their daily life, is even used as an indicator in research to assess the impact of data (e.g. IDC (2017)[120]). As access to more data provides more potentially useful insights, using more open data can help not only governmental bodies, but also individuals and businesses, to make better decisions based on more information available. Open data thus has a potential economic value that significantly exceeds its strict public sector utility. Moreover, another aspect of the economic value of open data is the value it can create through re-use. The fact that new data is available that was not accessible before stimulates creative minds to explore new business opportunities. By allowing re-use of governmental data, knowledge can be generated from analysing the available data which, in turn, can be applied in services and products based on or enriched by open data.

Existing studies

A growing body of literature studies the economic impact of open data on national and international level (see a selection of the literature in chapter 1). Yet, relatively few have thoroughly researched the causal relationship of open data re-use and value created. Notable examples are a report by McKinsey (2013)[121] and a report by Lateral Economics (2014)[122]. Both take an approach of researching several sectors in which open data can have an influence, both refer to use cases, and both try to extrapolate findings to calculate a global - or local - potential of value created. However, these and other researches emphasize the difficulty of quantifying the value created. In most cases, research reports the potential positive causal effects of open data re-use in a qualitative manner and - if any - only quantify it to a limited extent. In addition, a nuance that is often missing is a clear distinction between actual open data use cases and use of to-be opened data. Some estimations might be unrealistic or at least overestimated as they are based on data that is not - and might be unlikely to ever become - open data. Though there are some limitations to these researches - as there will undoubtedly be some limitations in this report – their main approach serves as an anchor for this research.

---

120 IDC (2017) available at: https://a2528ba5-a-c3c32646-s-sites.googlegroups.com/a/open-evidence.com/download/repository/SMART20130063_Final%20Report_030417_2.pdf?attachauth=ANoY7crL55TcXaDjVhYlzdD2jBNz6-UaN35LRS8sfZ9bqquqBazhj8Xk5FFQaus6z1xdbaviMZQ0WbjkhstO9Wpv7a13pztoDD9FrCANHECqEHeODrZgkvAvEes5w0ZiwC5t8L0Ovpk_7CHJVVPmUZWeUiVCYQoRGYk_b5ln0ZeTP3TH_A2sJhNeFtAg8X0ap3CqXfqHFmJGtgZkkEr0bcyM0pbRtH7UZ76oxlk89xXYB3FhdJ096eQqfhBG6Cx0TBGI9VCSKQRE&attredirects=0; http://datalandscape.eu/webinars/how-power-data-can-drive-european-union-economy-european-data-market-study-and-monitoring

121 McKinsey (2013) available at: https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/Open%20data%20Unlocking%20innovation%20and%20performance%20with%20liquid%20information/MGI_Open_data_FullReport_Oct2013.ashx

122 Lateral Economics (2014) available at: https://www.omidyar.com/sites/default/files/file_archive/insights/ON%20Report_061114_FNL.pdf

Methods

In a similar vein as these reports, this research aims to provide the reader with a thorough understanding of the causal chains through which open data can help create efficiency gains in several topics. Moreover, this report aims to quantify the potential efficiency gains as much as possible, e.g. as the potential number of lives saved, or the potential amount of time saved.

There are many societal gains from sharing open data and the products and services that are at least partially based on open data. These can reflect in (but are not limited to) efficiency gains related to:

- Saving lives in threatening situations, such as heart failures and traffic accidents, but also related to transmittable diseases.
- Saving time on various tasks, such as commuting via public transport or in traffic.
- Benefits for the environment, such as reduced air pollution or energy use, but also encouraging alternative energy sources, increasing biodiversity, and fostering more sustainable ways of living.
- Increase in knowledge transfer, for example by a better quality of research through access to more data, but also by improved translation services.

For each of these types of efficiency gains, the causal chains and corresponding quantifications will be described. By using sound and transparent logic, reliable data sources such as Eurostat, and insights from academic peer-reviewed articles we aim to safeguard the reliability of our calculations and underlying assumptions. These will be made tangible by use cases of actual open data, examples of to-be opened data, and examples that go beyond open data, e.g. combining open data with other types of data. The calculations are always an approximation and if assumptions are made these will always be clearly marked as being assumptions.

In the following sections, the beforementioned efficiency gains for society will be elaborated on further. Each section will be substantiated with examples (actual, to-be, and beyond open data) and will guide you through the calculations of potential gains.

Saving lives with open data

290 - 400 thousand lives saved

due to better allocation of resources to combat malaria globally*

54 - 202 thousand lives saved

thanks to emergency services arriving at the scene of an incident 1 minute faster

14 - 22.3 thousand lives saved

thanks to first responders administering CPR before the arrival of emergency services

* Potential lives saved due to enhanced malaria treatment are global numbers.

The above estimations are potential numbers for EU28(+) or EU27(+) depending on availability of data. For details on calculations and assumptions see corresponding sections.

EUROPEAN DATA PORTAL

5.2 Saving lives

One of the more impactful effects that open data can have through the development of new products and services is saving lives. Lives are not only saved within hospital walls and with the help of personal medical records. Openly shared data can also play a role in practices that help to save people's lives. In the following section we discuss how open data can contribute to reducing infectious diseases globally, improving the efficiency of European emergency services, and enabling better care in the time between an incident and the arrival of an ambulance.

5.2.1 Transmittable disease outbreaks

Diseases that are transmittable by mosquito bites such as malaria, dengue fever, and yellow fever not only cause people to suffer the disease's symptoms. In certain cases, these infectious diseases can be fatal. In 2017, an estimated 219 million cases of malaria occurred worldwide and there were an estimated 435 000 deaths from malaria globally.[123] Open data and location analytics are used to combat malaria by an initiative called "Visualize No Malaria"[124] - an online tool created by Zambia's Ministry of Health in collaboration with health organisation PATH and the Tableau Foundation.

2019 Geospatial World Excellence Award winner - Visualize No Malaria[125] uses open data and location analytics to work towards eliminating malaria. Thousands of volunteers have worked to map hundreds of thousands of square kilometres of the malaria-affected world, transforming how Province and District Health Care Professionals report, visualise, and act on their data, leading to more effective distribution of resources and identification of areas of high need. The improved collection and use of data have contributed to an 85% reduction in reported malaria cases and a 92% reduction in malaria-related deaths across Zambia's Southern Province from 2014-2017 - affecting the equivalent of about 1.8 million people. Similarly, a 60% reduction in malaria cases is reported by the Visualize No Malaria tool in Senegal. Maps created by the service feature data from Mapbox and OpenStreetMap and their data partners.

If the results of the pilot in Zambia and Senegal could be applied in other areas affected by mosquito transmittable illness, it can be assumed that, with the help of the Visualise No Malaria tool, 60 to 85% fewer people will get infected as resources related to prevention are allocated more effectively. This would lead to a global decrease of 131.4 - 186.2 million cases of malaria.[126] Additionally, the people that do get infected by the disease will get better care, leading to a total decrease of fatalities of 67-92%[127], i.e., 291 450 - 400 200 potential lives saved globally.

[123] World Health Organization Malaria Report (2019) available at: https://apps.who.int/iris/bitstream/handle/10665/275867/9789241565653-eng.pdf?ua=1
[124] Tableau on "Visualize No Malaria" (2019) available at: https://www.tableau.com/solutions/customer/path-and-zambian-ministry-health-use-data-track-down-last-malaria-parasites
[125] Mapbox on "Visualize No Malaria" (n.d.) available at: https://www.mapbox.com/showcase/path
[126] Global 219 million cases * 60% to 85% = 131.4 to 186.2 million
[127] Note: As Senegal showcased a 60% decrease where Zambia showed 85% decrease in reported cases, for the fatalities the same difference is used (25%) resulting in the interval of 67%-92%

Globally, open data can potentially help save 290 - 400 thousand lives due to better allocation of resources to combat malaria.

Luckily, there are only a minimal amount of cases of malaria in Europe of which around 99% are travel related. Sporadic (5 - 10 cases a year) locally acquired malaria cases have been reported in the EU that were related to either transmission by a local mosquito infected by a returning traveller or by an infected mosquito transported by aircraft from malaria-endemic countries.[128] The technique developed to reduce malaria with the help of open data, not only makes it safer for people to travel, but in case of an outbreak, it will enable faster and better response.

5.2.2 Emergency services

Each year, approximately 293 million emergency calls are made across Europe of which 141 million are to the Single European Emergency number "112".[129] Open data can help emergency services reach the location of an incident faster, which, in turn, helps to save more lives. To optimise the emergency services, several types of open data can be used, for example:

- Open data on emergency numbers
- Open maps data
- Open real-time traffic data

In this section, several ways in which open data can optimise the time to location are discussed. Firstly, due to open maps data and open government data people can more easily find the correct local emergency number. Secondly, when calling an emergency number, open maps data can help to optimise identification of the caller's location. Third, with real-time open traffic data, the routing of an ambulance to the scene of an incident can be quicker.

Finding the correct emergency number faster

Almost everyone knows the national emergency number of their home country. Yet, imagine you are on holiday or travelling internationally, would you readily have the local emergency number at hand? Making national and local emergency numbers openly available can help people to get in contact with local emergency services faster. Based on your current location, applications such as "Emergency Call"[130] or "TripWhistle Global SOS"[131] can instantly redirect you to the local emergency number when travelling. Thereby, precious time is saved when in need of emergency services.

Beyond open data: Identifying the caller's location more accurately

Moreover, establishing your exact location can be difficult - even more so when you are not in your home environment. Open data could potentially be combined with other data sources to help on this aspect. Generally, a lack of availability of caller location occurs in less than 10% of the calls, yet in some countries there are higher rates of failure to provide caller location such as e.g. Belgium (33%) and Latvia (25%).[132] If open maps

[128] Eurpean Center for Disease Control (2018) available at: https://www.ecdc.europa.eu/en/news-events/new-map-shows-presence-anopheles-maculipennis-sl-mosquitoes-europe
[129] European Commission 112 Implementation Report (2019) available at: https://ec.europa.eu/digital-single-market/en/news/2018-report-implementation-european-emergency-number-112
[130] Emergency Call app (n.d.) available at: https://play.google.com/store/apps/details?id=com.collette.emergencycall&hl=en_US
[131] TripWhistle Global SOS (n.d.) available at: https://apps.apple.com/us/app/tripwhistle-global-sos-international-emergency-phone/id839654331
[132] European Commission 112 Implementation Report (2019) available at: https://ec.europa.eu/digital-single-market/en/news/2018-report-implementation-european-emergency-number-112

data is used to help determine the location of a person placing an emergency call, it can help save lives. There are applications that accommodate direct sharing - not openly - of your current location with a tap of a button (e.g. 112BE in Belgium[133] and the still-in-progress Pan-European Mobile Emergency App[134]). Additionally, phones with Advanced Mobile Location (see box below) can automatically send the caller's location during an emergency call.

When an emergency call is made with a smartphone that is **Advanced Mobile Location (AML)**[135] **enabled**, the phone automatically activates its location capability (GNSS – provided by EU's Galileo satellite system – or Wi-Fi) for 20 seconds to establish its position and sends this information via a text message to the emergency services. The radius of the service is 50 meters or less for 85% of the calls. This is a life-saving difference when compared with traditional location through the identification of one's cellular radio network cell that can have a radius of tens of kilometres in rural areas.[136]

Thanks to the location accuracy improvements brought by AML, it is estimated in the 2017 HELP112 report by the European Commission that an average of 30 seconds can be saved on every mobile emergency call made in EU, and more than 1.5 minutes on average for calls placed in rural environments. In the same report, the AML implementation in all mobile phones in Europe is estimated to potentially save 800 extra lives every year thanks to the improvement in emergency caller location solutions.[137]

Arriving faster to the scene of an incident

Due to a faster response time of ambulances getting to the scene of an incident, lives can be saved. Ambulances must be at the place of an incident on average within 7-8 minutes and must take no longer than 15 minutes in life-threatening cases (NL[138] & UK[139]). If ambulances need to be dispatched, open data can be beneficial to optimise the time to location, as the routing to the scene of an incident can be faster if informed by real-time traffic data.

Real-time open traffic data enables users of navigation applications to be navigated along the fastest route from A to B. This is especially beneficial if the route they would take otherwise - most likely the *shortest* route - is congested by traffic.

In the Netherlands, there are on average 53 life-threatening ambulance rides per 1000 inhabitants in 2017.[140] In the United Kingdom, on average 99 life-threatening ambulance rides per 1000 inhabitants are made.[141]

133 112BE (n.d.) available at: https://www.112.be/nl/app
134 PEMEA (n.d.) available at: https://eena.org/apps/
135 EENA Help112 Report (2017) available at: https://eena.org/help-112-findings/
136 EENA Help112 Report (2017) available at: https://eena.org/help-112-findings/
137 EENA Help112 Report (2017) available at: https://eena.org/help-112-findings/
138 Ambulances in the Netherlands (2017) available at: https://www.ambulancezorg.nl/sectorkompas/ambulancezorgverlening-in-2017
139 Ambulances in the United Kingdom; NHS Ambulance Services Report (2017) available at: https://www.nao.org.uk/wp-content/uploads/2017/01/NHS-Ambulance-Services.pdf Note: 51% * 195 ambulance rides per 1000 inhabitants
140 Ambulances in the Netherlands (2017) available at: https://www.ambulancezorg.nl/sectorkompas/ambulancezorgverlening-in-2017
141 Ambulances in the United Kingdom; NHS Ambulance Services Report (2017) available at: https://www.nao.org.uk/wp-content/uploads/2017/01/NHS-Ambulance-Services.pdf Note: 51% * 195 ambulance rides per 1000 inhabitants

Assuming that the same numbers can be applied to Europe, 53-99 life-threatening ambulance rides per 1000 inhabitants, which given the EU28+ 527.1 million inhabitants[142], would result in approximately 29 - 52 million ambulance rides made in Europe.

A study examining emergency incidents in the Salt Lake City area (US) found that on average, a one-minute decrease in ambulance response times reduced the likelihood of death 90 days after an incident happened from 6% to 5%, i.e., a 17% decrease in the total number of deaths.[143]

> Approximately 54 – 202 thousand lives can be saved in the EU because emergency services arrive at the scene of an incident 1 minute faster.

In the 12 most congested metropolitan areas of Europe, one minute can be saved using real-time traffic data in 19.4% of rides during the day and 38.9% of rides in peak hours[144]. If we assume that the same numbers can be applied for Europe as a whole, using real-time traffic data based on open data helps ambulances to realise a one-minute earlier arrival of emergency services in approximately 5.4 - 10.1 million rides in Europe.[145]

A 1% decrease of the likelihood of death of people served by European ambulance rides (assuming the results of the Salt Lake City study can be applied to Europe) would result in, approximately 54 - 101 thousand potential lives saved thanks to open data.

5.2.3 First responders

Not only does the earlier arrival of emergency services have the potential to save lives with the help of open data. There is a crucial gap between the time of an incident and the arrival of emergency personnel. With the help of open maps data-based apps, first responders in the neighbourhood can be guided to the location of an accident (e.g. traffic accident, cardiac arrest, or stroke) to administer first aid more quickly than the time it takes for the ambulance to arrive. If needed, they could also be guided to the nearest automatic external defibrillators (AEDs). The example of cardiac arrest is elaborated below.

Providing first aid to people having a cardiac arrest

In Europe, on average 84 per 100 000 of European citizens[146] have a cardiac arrest outside the hospital, which amounts to approximately 443 thousand European citizens annually. Every minute without cardiopulmonary resuscitation (CPR) and defibrillation reduces the chance of survival by 10%[147]. If effective CPR is administered within three to five minutes and can double or triple a victim's chance of survival[148]. It is estimated that if CPR is administered 1 minute earlier, it could save approximately 7 000 lives in Europe[149].

---

142 EU28 (513.5mil) available at: https://ec.europa.eu/eurostat/statistics-explained/index.php/Population_and_population_change_statistics + EFTA (14.2mil) (Iceland 354,441, Norway 5,332,232, Switzerland 8,542,304, Liechtenstein no data) available at: https://appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do.

143 Wilde (2013) Study of 73 706 emergency incidents in U.S. available at: http://www.emdac.org/docs/Wilde_EMS%20Response%20Times%20&%20Outcomes_Health%20Econ_2013.pdf).

144 Capgemini research on the 12 most congested cities of Europe, considering 3 routes per city that take on average between 6.5-18 minutes.

145 Assuming ambulance rides are evenly distributed in a day. The average 19.4% is calculated by considering 8 hours (06.00-10.00 & 16.00-20.00) as peak hours (38.9%), 8 hours (10.00-16.00 & 20.00-22.00) as off peak (19.4%) and 8 hours (22.00-06.00) as night time (0%).

146 Gräsner, Lefering & Koster (2016). Study of 10682 out of hospital cardiac arrests in 27 European countries. Available at: https://www.sciencedirect.com/science/article/pii/S0300957216300995#bib0155

147 American Heart Association (n.d.) available at: https://www.zoll.com/-/media/uploadedfiles/public_site/core_technologies/real_cpr_help/cpr-fakten-pdf

148 American Heart Association (n.d.) available at: https://cpr.heart.org/en/resources/what-is-cpr

149 European Data Portal (2015) available at: https://www.europeandataportal.eu/sites/default/files/edp_creating_value_through_open_data_0.pdf

A study by Gräsner et al. (2016) on out-of-hospital cardiac arrests (OHCAs) across Europe found that CPR was started by a bystander 31.7% of cases. CPR is started by emergency medical services in 35.2%, and CPR is not attempted in 33.1% of cases.[150]

Another study by Bürger et al. (2018) of OHCAs in Germany found that there is a substantial difference in survival rates of people who do receive CPR from bystanders before the emergency services arrive (13.1% - 22% discharged alive) as opposed to those who do not (7.3% - 12.9% discharged alive).[151]

Let us assume that the survival rates of the German study can be applied to Europe in general. Of the 443 thousand European citizens that have a cardiac arrest, approximately 140 thousand people receive CPR from bystanders, 156 thousand people receive CPR only when the ambulance arrives, and 147 thousand people do not receive CPR at all. This, in turn, leads to an amount of people surviving of between 18.4 - 30.9 thousand in case of bystander CPR and 11.4 - 20.1 thousand in case of first CPR received by emergency services (see table on next page).

Assuming the amount of cardiac arrests and survival rates will remain the same, open data can be used to save lives by enabling more people to receive bystander CPR. There are several examples that are already in use, such as HartslagNu (NL), GoodSAM (UK), and PulsePoint (US). Additionally, examples such as Shock (Vienna) further help first responders with a map of available automatic external defibrillators (AEDs).

HartslagNu (NL) +50.000 downloads
GoodSAM (UK) +50.000 downloads
PulsePoint (US) +500.000 downloads

GoodSAM[152] (UK) is an application that uses GPRS technology to alert first responders that are close to the emergency location. People who downloaded the application can press a button to call for help while their location is being identified. There are 5 000 medical emergencies each day in London alone, but also thousands of individuals that can provide medical assistance while waiting for an ambulance to arrive.

Once emergency dispatchers get a call about a suspected cardiac arrest in a public place, they activate an alert to **PulsePoint**[153] (US) application users at the same time they send out first responders. Users are notified if they are within a certain distance and if the victim is in a public location. The application also directs citizen responders to the place where they can find the nearest publicly accessible defibrillator, a device that sends an electric shot to the heart to try to restore its normal rhythm.

150 Gräsner, Lefering & Koster (2016). Study of 10682 out of hospital cardiac arrests in 27 European countries. Available at: https://www.sciencedirect.com/science/article/pii/S0300957216300995#bib0155
151 Bürger et al. (2018). Study of 10853 out of hospital cardiac arrests in Germany 2010-2016. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6156551/
152 GoodSAM (n.d.) available at: https://www.goodsamapp.org/
153 Pulsepoint (n.d.) available at: https://www.pulsepoint.org/

In Europe, approximately 69.4% of out of hospital cardiac arrests occurred in a private residence and the remaining 31.6% of cases are assumed to occur in a public place.[154] So, at least this group can potentially be aided by open data applications guiding first responders to the scene of a cardiac arrest. If 31.6% of the groups currently not receiving bystander CPR could potentially get bystander CPR because of applications developed based on open data (i.e., 31.6% of 156 = 49, and 31.6% of 147 = 46), the group receiving CPR by bystanders would grow from 140 thousand to 235 thousand across Europe (140 + 49 + 46). Of this group, 13.1% - 22% survives and as opposed to the current situation, this would result in 8.8 - 13.7 thousand potential lives saved. So, open data can be used to not only help people get bystander CPR faster, but also to increase the group of people receiving bystander CPR.

> Approximately 14 - 22.3 thousand lives can potentially be saved in the EU28+ because first responders administer CPR before the arrival of emergency services.

This number could potentially be even higher It could be that, for example, in about half of those instances in a private residence there is someone else present or the person him-/herself is able to alarm others and open the door for help. This group could then be helped by the initiatives above in addition to the 31.6% of cases that occur in a public place.

Table 14: Overview of current vs. potential number of survivors depending on receiving CPR or bystander CPR

	Current		Potential with open data-based applications	
	Number of people with OHCAs	Number of people surviving OHCAs	Number of people with OHCAs	Number of people surviving OHCAs
First CPR by bystanders (13.1% - 22%*)	140	18.4 – 30.9	235 (140 + 49 + 46)	30.8 – 51.7
First CPR by emergency services (7.3% - 12.9%*)	156	11.4 – 20.1	107 (156-49)	7.8 – 13.8
No CPR (0%)	147	0	101 (147-46)	0
Total	443	29.8 – 51	443	38.6 – 65.5
Additional lives saved				8.8 – 13.7

* = survival rate interval based on the arrival time of emergency services
Note: numbers are in thousands.

[154] Gräsner, Lefering & Koster (2016). Study of 10682 out of hospital cardiac arrests in 27 European countries. Available at: https://www.sciencedirect.com/science/article/pii/S0300957216300995#bib0155

Saving time with open data

27 million hours saved

due to reduction of waiting time because of real-time notifications in case of delayed trains

500 – 730 million hours saved

due to reducing time stuck in traffic with real-time traffic navigation

The above estimations are potential numbers for EU28(+) or EU27(+) depending on availability of data. For details on calculations and assumptions see corresponding sections.

EUROPEAN DATA PORTAL

5.3 Saving time

In addition to saving lives, time too can be saved by the efficiency gains that are achieved with the use of open data. Doing something more efficiently can refer to doing more in the same amount of time or doing the same task or activity in a shorter amount of time. In this section, the focus is one specific kind of timesaving, namely travelling. Most of the working population in Europe does not work from home, as only 5% usually work from home.[155] Saving time on travelling is thus beneficial for a large group of people. We will discuss time saved in public transport and time saved in traffic with the help of open data.

5.3.1 Public transport

Open data on public transport routes, departure times, delays, etc., can help to optimise public transport in Europe. Optimising public transport with open data can apply to the public transport system within cities (e.g. bus, metro, tram, etc.), but also to the 9 billion passengers who every year use national railway networks, the 1 billion passengers travelling by air, and the 415 million maritime transport passengers.[156]

For some time now, schedules of public transport have been made available as open data. Moreover, people travelling by train, for example, are able to plan their journeys better with apps that rely also on open data to provide real-time travel information and advice on how to adjust their route. This helps them to save time because they instantly have accurate and reliable information to:

- anticipate a potential delay and thereby optimise the duration of the journey - reducing waiting time as people use an app to plan their walk to a train station at the exact right time, whether the train is on time or delayed, and
- adjust their route choices - reducing travel time as an alternative route can be quicker than a normal route that currently has issues or delays

The time saved because of anticipating a potential delay with the use of real-time public transport data will be elaborated on below.

With **Real Time Ireland App**[157], users receive real-time information on Irish public transport. The application allows users to navigate through stops and routes and to search by route or by stop number. In addition, users can set "favourite" routes, which enables quick access to real-time travel data regarding them. The application reduces waiting times at stations and stops because users can activate alerts for approaching buses or trains.

Waiting time in public transport

In academic literature (see Brakewood & Watkins (2018) for an overview) it is demonstrated that real-time information provided via personal devices decreases both the actual waiting time and the time people feel they have been waiting, i.e., perceived waiting time. Real-time information can benefit passengers before they leave their point of origin (reducing actual waiting times) and at the (bus-)stop or station

[155] Eurostat (2018) available at: https://ec.europa.eu/eurostat/web/products-eurostat-news/-/DDN-20180620-1
[156] Rail, air, and maritime passengers EU28+. Eurostat (2019) available at: https://ec.europa.eu/eurostat/statistics-explained/index.php/Passenger_transport_statistics#Rail_passengers
[157] Real Time Ireland App (n.d.) available at: https://play.google.com/store/apps/details?id=com.osds.rtpi

(decreasing perceived waiting times).[158] Especially the perceived waiting times can affect satisfaction with public transport services, as people generally prefer to not feel uncertain about their journey.[159] A study by Watkins et al. (2011) in Seattle (US), showed actual waiting time reductions of approximately 2 minutes and 30% less perceived waiting times due to real-time information available.[160] Another study in a rural area in Scotland by Papangelis et al. (2016) showed an even greater decrease of approximately 60% perceived waiting time with real-time public transport information.[161]

Time saved in public transport

In the EU28+, more than 9 billion passengers use the national railway services each year.[162] Together they account for approximately 450 billion kilometres of train rides.[163] The punctuality rates in all EU countries range from 80.1% (Bulgaria) to 99.3% (Estonia).[164] The remaining train rides are delayed by more than 5 minutes (or might not ride at all – but that is not considered in the following calculation). Based on the numbers above, the total number of people on delayed train rides in Europe[165] is estimated at approximately 820 million people each year.[166]

> Open data can potentially save approximately 27 million hours annually for train users in Europe.

Let us imagine that in an ideal open data-empowered world all these people are aware of the delay before they leave home because they can find that out thanks to an app on their phone. Applying the results of the Watkins et al. (2011) study, this would on average lead to an actual waiting time reduction of 2 minutes per person each time a train is delayed, resulting in a total of approximately 1.6 billion minutes (or 27 million hours) saved by all train users in Europe on a yearly basis. Moreover, the perceived reduction in waiting time can be even higher. Based on the findings by Watkins et al. (2011) and Papangelis et al. (2016), perceived wait time reduction ranges between 2.4 - 4.8 minutes[167], so 2.0 - 3.7 billion minutes (or 33.3 - 61.7 million hours) perceived waiting time saved.

The above example is merely focused on train use, but the same principle can apply to any other mode of public transports, thus potentially even more time can be saved in public transport with the help of open real-time data.

158 Brakewood & Watkins (2018) available at: https://www.tandfonline.com/doi/full/10.1080/01441647.2018.1472147?scroll=top&needAccess=true
159 Brakewood & Watkins (2018) available at: https://www.tandfonline.com/doi/full/10.1080/01441647.2018.1472147?scroll=top&needAccess=true
160 Watkins et al. (2011) available at: https://www.sciencedirect.com/science/article/pii/S0965856411001030
161 Papangelis et al. (2016) available at: https://www.tandfonline.com/doi/abs/10.1080/03081060.2015.1108085?journalCode=gtpt20
162 Rail passengers EU28+ https://ec.europa.eu/eurostat/statistics-explained/index.php/Passenger_transport_statistics#Rail_passengers
163 European Commission Report on monitoring development of the rail market (2019) available at: https://ec.europa.eu/transport/modes/rail/market/market_monitoring_en
164 European Commission Report on monitoring development of the rail market (2019) available at: https://ec.europa.eu/transport/modes/rail/market/market_monitoring_en
165 Excluding Switzerland, Norway, Greece and Spain due to missing data
166 Capgemini research. From total number of kilometres travelled by train by all train users in each country, the total number of train rides by all passengers in a country is estimated and based on the punctuality rates the % of delayed trains is calculated and applied to the total number of train rides by passengers.
167 The perceived time saved while informed by real-time information in the Papangelis et al. (2016) study is approx. 60% in rural areas which is applied to the numbers of the Watkins et al. (2011) study (i.e., 30% decrease is 2.4 minutes). So, if 30% decrease is 2.4 minutes, then 60% is 4.8 minutes decrease.

Transport for London (TfL)[168] makes data about its network, including the Tube, Docklands Light Railway, London Overground, buses, TfL Rail and tram, and cycle hire, available through APIs, downloads, and feeds. Businesses, academics, and developers partner with TfL and use this data to create new commercial and non-commercial products and services for customers.

Moreover, as applications built on open data can reduce waiting and travel times, not only time is saved, but people can be more satisfied with the experience of public transport. In turn, open data can potentially help increase the willingness of people to use public transport as a more environmentally friendly alternative to private transport.

5.3.2 Traffic

Open data on traffic can help to optimise the flow of traffic on roads in Europe. When driving from home to work and back or for any other reason, we would all like to be at the destination as quickly and smoothly as possible. However, as it is known, this is not always the case. This is mostly due to traffic jams that are a mere consequence of many people wanting to get to work around the same time. Yet, it could also be because of traffic-related to concerts and sports matches, or if unexpected events happen such as a car breakdown or an accident. In any case, the usual route will get congested and you might find yourself stuck in traffic. Open real-time traffic data can help find alternative routes that are less congested. Taking the alternative route does not only help you get to your destination faster: it might even help to dissolve the traffic jam faster as cars re-distribute across roads when using alternative routes.

In Norway, the **Trafikkflyt**[169] app provides traffic density indications to help decrease time spent in road congestion and improve the flow of traffic. Information from the National Public Road Administrations is used to create maps of the traffic density. Events held are listed and pointed on a map, so that people can avoid crowded areas.

OpenTraffic[170] is a global data platform to process the position of vehicles and smartphones anonymously, making them available into real-time and historical traffic statistics. OpenTraffic aggregates global, freely available traffic speed data linked to OpenStreetMap. It is built using fully open-source software, with involvement from Conveyal, World Bank, Mapzen, and Mapbox.

Time spent in traffic jams

Research by TomTom shows that in 2017, the average yearly number of hours spent by drivers in road congestion in Europe ranges from 18 hours (Finland) to almost 46 hours (United Kingdom).[171] These are average numbers

[168] Transport for London (n.d.) available at: https://tfl.gov.uk/
[169] Trafikkflyt (n.d.) available at: http://www.trafikkflyt.no/
[170] OpenTraffic (n.d.) available at: http://opentraffic.io/
[171] TomTom (2017) available at: https://ec.europa.eu/transport/facts-fundings/scoreboard/compare/energy-union-innovation/road-congestion_en

by the average driver across the country. However, these numbers are much higher in large cities, where there is generally more traffic. In the European cities that are among the 200 most congested cities worldwide, the average time wasted in traffic ranges from 33 hours (Cordoba, Spain) to 254 hours (Rome, Italy) annually per driver[172] and is on average 129.5 hours in all European cities considered in this study. There is thus much more time to be saved in the European cities as opposed to the more rural areas.

Time saved in traffic

In the 12 most congested metropolitan areas of Europe, using real-time open traffic data can help decrease the time spent in traffic by approximately 3.8% during daytime with 5.5% during peak hours.[173] If we assume that the same numbers can be applied to all cities in Europe, using real-time traffic data based on open data helps to decrease time in traffic by 3.8 - 5.5% across European cities. In Rome, for example, this would mean that using real-time open traffic data can have a potential effect of 9.7 - 14 hours saved per driver. In European cities on average, drivers could save 4.9 - 7.1 hours annually.

> Open real-time traffic data can help decrease the time a driver spends in traffic by approximately 4.9 - 7.1 hours annually in European cities.

Of the total 513.5 million inhabitants in the EU28[174], the total number of people in employment reached 220.7 million in the EU28 in 2015.[175] Approximately 72% of the European population lives in an urban area[176], in which time spent in traffic is highest. So, there are roughly 159 million working Europeans living in urban areas who could benefit the most of potential time saved in traffic.

Between 60% to 70% of commuting from and to work is done by car in Europe.[177] So, let us assume that on average 65% of the working European population living in urban areas commutes to work by car, i.e., 103 million people. This group can save the previously mentioned 4.9 - 7.1 hours annually with the help of open data based real-time traffic applications. In total, this results in approximately 500 - 730 million hours potentially saved each year by European drivers commuting to and from work in urban areas.

> Open traffic data can potentially save 500 – 730 million hours for Europeans commuting by car.

172 INTRIX Global Traffic Scorecard (2018) available at: http://inrix.com/scorecard/
173 Capgemini research on the 12 most congested cities of Europe, considering 3 routes per city that take on average between 6.5-18 minutes.
174 EU28 (513.5mil) available at: https://ec.europa.eu/eurostat/statistics-explained/index.php/Population_and_population_change_statistics + EFTA (14.2mil) (Iceland 354,441, Norway 5,332,232, Switzerland 8,542,304, Liechtenstein no data) available at: https://appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do.
175 Eurostat (2015) Number of persons employed in EU28. Available at: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Archive:Statistics_on_commuting_patterns_at_regional_level
176 Eurostat (2017) Population in rural and urban areas in EU28. Available at: https://ec.europa.eu/eurostat/statistics-explained/index.php/Statistics_on_rural_areas_in_the_EU
177 Summary of studies in France, Germany, United Kingdom, the Netherlands, and Belgium. https://www.fleeteurope.com/en/smart-mobility/europe/features/car-remains-primary-means-commuting-western-europe?a=SBL09&t%5B0%5D=Traffic&t%5B1%5D=study&t%5B2%5D=France&curl=1 and Study in Poland http://www.research-pmr.com/userfiles/file/wp/wp_37_894_2013-01-23%20-%20The%20car%20is%20the%20most%20popular%20means%20of%20commuting%20to%20work%20-%2001.pdf

Beyond open data:

In addition to real-time traffic data, open crowdsourced traffic data generated by a community of users can also be used. As the world's largest community-based GPS navigation application, **Waze**[178] functions like a crowdsourced social network that provides directions using real-time traffic data and community feedback. With Waze, users can alert others to accidents, alternative routes, road closures, police officer sightings, for example. In 2017, Opendatasoft announced a partnership with Waze make it easier for cities and other local governments to launch smart transportation initiatives as they can quickly and seamlessly collect, harmonise and share data through the Waze Connected Citizens Program, a free two-way data share of publicly available traffic information.[179] Combining real-time traffic data with other kinds of (open) data, e.g. crowdsourced data in the example of Waze, can potentially even further increase the amount of time saved.

The beforementioned numbers are based on open data that is available already today, however, the potential for further improvement is even more significant. For example, in August 2019 the UK government announced plans to make data on congestion, repair works, and any other scheduled disruptions to Britain's road network openly available. The aim is to enable third parties to develop apps that can provide not only real-time traffic information, but also ahead-of-time information to drivers thanks to predictive modelling and AI.[180] When open data based applications for traffic reduction are even further optimised with these types of data, and across Europe, the potential time and costs saved can be even higher than described above.

[178] Waze (n.d.) available at: https://www.waze.com/en
[179] Opendatasoft (2017) available at: https://www.opendatasoft.com/blog/2017/06/28/opendatasoft-partners-waze-deliver-unprecedented-transportation-oriented-smart-cities-initiative
[180] Department for Transport UK (2019) available at: https://www.gov.uk/government/news/artificial-intelligence-to-end-future-holiday-jams-caused-by-roadworks

Helping the environment with open data

5.8 million tonnes of oil equivalent saved

due to reduced energy consumption by households

353 extra terawatt hours of solar energy

thanks to efficient use of solar panels on rooftops

Supports EU sustainability goals

by increasing the share of renewables in gross final energy consumption at 32% for 2030

The above estimations are potential numbers for EU28(+) or EU27(+) depending on availability of data. For details on calculations and assumptions see corresponding sections.

EUROPEAN DATA PORTAL

5.4 Helping the environment

Climate change is a topic of growing concern. The consequences of the greenhouse effect are described in numerous research studies and policy papers including the United Nations Framework Convention on Climate Change treaty of 2015.[181] Under the subsequent Paris Agreement[182], each signatory country must determine, plan, and regularly report on the contribution that it undertakes to mitigate climate change. Governments are thus searching for new ways of reducing CO2 emission, reducing waste, and also to activate and enable individuals to take an active part in this goal. The following sections focus on how open data can contribute to reducing the total energy consumption in Europe, increasing the use of sustainable energy sources, protecting biodiversity, and increasing awareness on CO2 emissions.

5.4.1 Reducing energy consumption

One of the ways in which open data can help attain the CO2 emission goal is by informing initiatives to encourage people to reduce their energy usage. In the EU28 countries, residential households account for 27.2% of the total energy consumption of 1 060 million tonnes of oil equivalent (Mtoe), i.e., approximately 288.3 Mtoe, in 2017.[183] Open data-based applications can serve as a tool to help households reduce their energy consumption by providing them with reports and suggestions on how to decrease energy usage. In these reports, open aggregated energy consumption data is combined with household energy data (not open). For example, by comparing the energy use of households to - anonymised data of - other households that are equivalent in terms of number of family members and consumption patterns, people can be socially motivated to pollute less by reducing their consumption while at the same time cutting on their energy bill costs. These reports thus aim to serve as a behavioural intervention that affects energy consumption habits.

In Denmark, **Modstroem Energiberegner** (formerly Husets Web)[184], provides intelligent and dynamic calculations of energy saving measures that can help homeowners save energy and money. The calculator uses information on Danish building architecture, 200 years of building statistics, and information about heating systems. This information is combined with data on the specific property which is drawn from public databases. In addition, the homeowner can add her own data about the property and energy consumption. The calculator produces an individual "Energy Report" which offers opportunities for improvement.

[181] UNFCCC (n.d.) available at: https://unfccc.int/process-and-meetings/the-convention/what-is-the-united-nations-framework-convention-on-climate-change
[182] Paris Agreement (2015) available at: https://unfccc.int/process-and-meetings/the-paris-agreement/what-is-the-paris-agreement
[183] Eurostat Energy Consumption (2019) available at: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Energy_consumption_in_households#Energy_products_used_in_the_residential_sector
[184] Modstroem Energiberegner (n.d.) available at: https://www.modstroem.dk/energiberegner/

Opower[185], now a part of Oracle, says to save $2 billion (€1,8 billion[186]) in utility bill savings globally, by reducing energy consumption. Opower provides people with information on their own energy consumption, but also how their consumption compares to similar households. Opower relies on open US government data such as the Residential Energy Consumption Survey (RECS) to understand how households are using energy around the country. Opower utilises this data, along data from the US Census Bureau on the mix of gas and electric heating sources in a given region, to create location-specific baselines to use when analysing an individual's home energy consumption. This, in turn, helps Opower deliver personalised information to individual customers.[187]

A research by Alcott and Rogers (2014) on the effectiveness of home energy reports on energy reduction in the US found that the open data-based Opower program reduced energy consumption among treated households by 1.4%-3.3%. The average reduction was 2.0% or 0.62 kWh per day.[188] For illustration, 1 kWh can enable you to cook in a 2 000-watt oven for half an hour or watch around three hours of television on a plasma TV[189].

Open data-based tools could potentially save 5.8 million tonnes of oil equivalent by helping to reduce household energy consumption.

If we assume that the same average energy reduction of 2% holds true for equivalent open data-based applications that could be used in Europe, the potential energy saved by households with the help of open data would be approximately 5.8 million tonnes of oil equivalent each year. Moreover, the beforementioned energy reports could be applied to not only individual households, but also to organisations and public buildings, potentially leading to even higher energy savings.

5.4.2 Increasing sustainable energy use

In addition to using less energy, open data could also play a role in making it more attractive or making it easier for people to switch to alternative sources of energy. The 2018 EU directive on the promotion of the use of energy from renewable sources sets the target for the share of renewables in gross final energy consumption at 32% for 2030.[190]

To help achieve this target, publicly available data on houses, utility usage, and location-based weather conditions can be used to provide tailor-made advice to people looking to make their house more energy-neutral, such as using solar panels or a solar water heater. In addition, the cost of an energy improvement solution can be directly estimated to interested homeowners, enabling them to instantly evaluate costs, potential benefit and time needed to make a return on the investment.

185 Opower (n.d.) available at: https://www.oracle.com/industries/utilities/products/what-is-opower.html
186 Exchange rates as of 26 January 2020
187 Energy.gov (2013) available at: https://www.energy.gov/data/articles/who-uses-open-data
188 Alcott and Rogers (2014) available at: https://www.povertyactionlab.org/evaluation/opower-evaluating-impact-home-energy-reports-energy-conservation-united-states
189 Ovo Energy 1 kWh usage (n.d.) available at: https://www.ovoenergy.com/guides/energy-guides/what-is-a-kwh-kw-and-kwh-explained.html
190 Council of European Union Directive 2018/2001 of the European Parliament and of the Council on the promotion of the use of energy from renewable sources (2018) available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1549740096222&uri=CELEX:32018L2001

GreenHome[191] is a Dutch website that offers independent advice on how to create a more sustainable house. GreenHome uses publicly available data from different sources, such as the Dutch cadastre, maps data, Actueel Hoogtebestand (up-to-date information on heights as opposed to the water level (NAP, Normaal Amsterdams Peil)) and data on utility usage. This data is used to train an algorithm that provides users with custom-made information to achieve an energy-neutral status. This information includes, for example, installation costs of solar, heating, insulation and other energy-saving measures that are relevant to their specific building.

Sunenergia[192], a Finnish company, provides information about the ideal area to cover in solar panels for optimal results, and where they should be placed. Sunenergia uses open weather data and location-based solar radiation data. It is an easy-to-use web-based decision-making tool for assessing the feasibility, viability and profitability of a solar energy system on any specific rooftop, using the slope of the roof and the exact amount of sunlight and shade it receives during any given hour of the year.

One of the ways in which the gross energy consumption can be sourced from more renewable energy is through the increased use of solar panels for electricity production. The Bloomberg 2018 New Energy Outlook forecasts a slight increase in the total electricity demand in the EU28+ from 3 454 TWh in 2017 to 3 566 TWh in 2030.[193] In 2017, the photovoltaic solar energy contribution in the EU28 to the electricity demand was just 114 TWh[194] or 9.8 million tonnes of oil equivalent[195], less than 4% of the forecast demand.

A recent study by Bodis et al. (2019)[196] researched the extent to which buildings in the EU28 countries can provide the space for significantly increasing the use of solar panels. Using data from the European Settlement Map and European Urban Atlas, for example, as well as Cadastre data on buildings and solar irradiance data, the total available technical potential is calculated for each EU28 country. This is the expected annual electricity output if 100% of the suitable rooftop systems are developed, independently of the cost. In addition, the potential for cost-competitive rooftop solar systems that produce electricity at a lower cost than the retail electricity prices in each country, i.e., the economic potential, is calculated.

Bodis et al. (2019) find that in the EU28 countries, there is a total rooftop area available for solar panels of 7935 km^2 with a technical potential of approximately 680 TWh per year. The economic potential is approximately 467 TWh in the EU28. Some countries show technical potential, but it is just not economically feasible yet for solar systems to compete on cost with traditional electricity production, e.g. technical versus economic potential in Poland (30 TWh; 0 TWh) or Hungary (18 TWh; 0 TWh). On the other hand, there are also countries that show both huge technical and economic potential such as

191 Greenhome (n.d.) available at: https://greenhome.nl/
192 Sunenergia (n.d.) available at: https://sunenergia.com/en/
193 Bloomberg (2018) available at: https://bnef.turtl.co/story/neo2018
194 Bodis et al. (2019) available at: https://www.sciencedirect.com/science/article/pii/S1364032119305179
195 Unit converter TWh to Mtoe (n.d.) available at: https://www.iea.org/statistics/resources/unitconverter/
196 Bodis et al. (2019) available at: https://www.sciencedirect.com/science/article/pii/S1364032119305179

France (125 TWh; 125 TWh) and Germany (104 TWh; 103 TWh). Moreover, the weather, geographical, and economical features of eight European countries make solar energy suitable to cover a significant (>20%) share of the electricity consumption at competitive costs.[197]

> Open data-based tools can aid in the potential increase of solar energy production of 353 TWh in Europe.

With the help of open data-based applications, European citizens, businesses and public administrations can be activated to make their homes, offices and buildings more energy-neutral. Specifically, the rooftops of buildings in the EU28 countries can be used to increase solar energy production from 114 to 467 TWh at current cost-competitive prices, an increase of more than 300%.

Furthermore, if due to technological advancements the production cost of energy by solar panels further decreases and/or the costs becomes lower than other energy sources due to European incentives, the full potential of rooftop solar systems can be achieved. In that case, open data can play a role in reaching the full technological potential of 680 TWh of solar power in Europe, aiding to an increased share of renewables for our energy consumption.

> Open data can help to reach the EU target for the share of renewables in gross final energy consumption at 32% for 2030.

5.4.3 Improving sustainability research

The potential of open data in transitioning to more sustainable energy sources is of course not limited to solar energy. Other alternative energy sources, such as wind or hydropower can be further researched and developed with the help of open data.

Exceedence[198] in the UK is a platform using wind and tidal open data from e.g. UK's Environmental Agency and the Marine Institute of Ireland to compare renewable technologies and projects to see which one provides the best return on investments. Information about robust and credible financial plans based around project implementation of wind, wave, tidal, offshore wind and combined technologies is provided.

Open data can not only be used to make the research on different types of alternative energy sources more insightful, it can even be used to make the current state of different types of research more transparent with open data-based applications such as Exceedence. With the help of these types of applications, investors can be better informed on how to allocate their investments in the development of renewable

[197] Bodis et al. (2019) available at: https://www.sciencedirect.com/science/article/pii/S1364032119305179
[198] Exceedence (n.d.) available at: https://exceedence.com/monetising-metocean-data-an-open-data-project/

energy technologies. This way, open data can have an even greater effect on the transition to renewable energy sources than described in the previous section if transparency can help convince more investors to invest in sustainable energy research.

5.4.4 Improving biodiversity

A major challenge for sustainable development is the preservation of biodiversity. Biodiversity not only relates to the extinction of certain species. Biodiversity boosts ecosystem productivity and ensures natural sustainability as healthy ecosystems can better withstand and recover from a variety of disasters.[199] Businesses and organisations can face important ecological risks as a result of biodiversity impacts and dependencies, such as increased costs due to limited natural resources or disrupted business operations.[200] Loss of biodiversity can have dramatic effects, for example, mortality of bee colonies can have an economic effect of approximately €150 billion worldwide.[201]

On European and international level, initiatives have started to foster biodiversity, such as the European Natura 2000[202] and Habitats Directive as well as the Convention of Biological Diversity 2011-2020[203]. To conserve biodiversity, the distribution of natural resources must be assessed and integrated with information on habitat conditions. Open data is having a substantial impact on the ability to understand how biodiversity is being affected, through a vast amount of available species data in e.g. the Global Biodiversity Information Facility (GBIF)[204].

The Global Biodiversity Information Facility (GBIF)[205] is an international network and research infrastructure funded by the world's governments and aimed at providing anyone, anywhere, open access to data about all types of life on Earth. It provides data-holding institutions around the world with common standards and open-source tools that enable them to share information about where and when species have been recorded. This knowledge derives from many sources, including everything from museum datasets to geotagged smartphone photos shared by amateur naturalists in recent days.

In addition, open data helps in:

- the monitoring of biodiversity;
- the development of an improved understanding of biodiversity patterns;
- the assessment of biodiversity's vulnerability to climate change.

Open data can thus help in predicting the consequences of changes and planning measures to counteract biodiversity threats.

199 Biodiversity (n.d.) available at: https://www.conserve-energy-future.com/what-is-biodiversity.php
200 OECD Report on Biodiversity for G7 (2019) available at: https://www.oecd.org/environment/resources/biodiversity/Annexes-Biodiversity-Finance-and-the-Economic-and-Business-Case-for-Action.pdf
201 European Parliament https://www.europarl.europa.eu/news/en/press-room/20180123IPR92314/urgent-action-needed-to-protect-eu-bee-population-urge-meps
202 Natura 2000 https://ec.europa.eu/environment/nature/natura2000/
203 CBD https://www.cbd.int/
204 Global Biodiversity Information Facility https://www.gbif.org/
205 Global Biodiversity Information Facility https://www.gbif.org/

Biodiversity in agriculture

Biodiversity helps to regulate nutrient-cycling and biological pest control, and provides pollination, sustaining diverse crops and genetic stocks for plant breeding.[206]

Bees and other insects sustain biodiversity by providing pollination for the majority of crops and wild plants growing in Europe. Globally, the Food and Agriculture Organization of the United Nations (FAO) estimates that of the 100 crop species that provide 90% of food worldwide, 71 are pollinated by bees.[207] However, in some member states the number of bee colonies has declined by more than 50% over the past years. According to the French National Institute for Agricultural Research, the mortality of bees, if not tackled, could cost €150 billion worldwide.[208]

The Good Growth Plan[209] is an initiative to address the challenges of feeding a growing world population sustainably. One of the key pillars of this commitment is to help biodiversity flourish. A key strategy is managing less-productive farmland alongside fields and waterways to reintroduce local species, provide buffers for soil and water, and connect wildlife habitats. The Good Growth Plan has six different databases showing aggregated hectares of farmland that benefited from biodiversity conservation practices and measures their contribution of these practices. From 2014-2017 there are 229 projects in 37 countries implemented which enabled enhanced biodiversity on a total of 5.6 million hectares of farmland.[210]

Agroknow[211] focuses on knowledge management and research on knowledge-intensive technology innovations for agriculture, food, and biodiversity. Technology innovations are fostered and shared. Organisations and people can get support on how to address societal and environmental challenges using solutions that are informed and enhanced by high-quality data, as Agroknow has an extensive experience in building services for the agricultural community.

Biodiversity in forests

Research by Liang et al. (2016)[212] on the relationship between biodiversity and productivity in forests shows that a continued biodiversity loss would result in an accelerating decline in forest productivity worldwide. Forest productivity refers to the growth in m3 per hectare each year. On average, a 10% loss in biodiversity leads to a 3% loss in forest productivity. There is thus a positive and consistent relationship between tree diversity and ecosystem productivity. Liang et al. further estimate that the economic value of biodiversity in maintaining commercial forest productivity alone is $166 billion (€150 billion[213]) to $490

206 FAO Sustainable Agriculture for Biodiversity Report (2018) available at: http://www.fao.org/3/a-i6602e.pdf
207 European Food Safety Authority http://www.efsa.europa.eu/en/topics/topic/bee-health
208 European Parliament https://www.europarl.europa.eu/news/en/press-room/20180123IPR92314/urgent-action-needed-to-protect-eu-bee-population-urge-meps
209 The Good Growth Plan (n.d.) available at: http://opendata.syngenta.agroknow.com/the-good-growth-plan-progress-data
210 The Good Growth Plan Progress Report (2017) available at: https://www.syngenta.com/~/media/Files/S/Syngenta/2018/Progres-Report-2017-EN-The-Good-Growth-Plan.pdf
211 Agroknow (n.d.) available at: https://www.agroknow.com/
212 Liang et al. (2016) available at: https://science.sciencemag.org/content/354/6309/aaf8957
213 Exchange rates as of 26 January 2020

billion (€444 billion[214]) per year.[215] With the help of open biodiversity data, research can further emphasize the importance of biodiversity to policymakers which, in turn, can take action to conserve biodiversity in forests.

Biodiversity in the oceans

In addition to land-based biodiversity, marine biodiversity is crucial to consider as well. According to the 2016 report of the Food and Agriculture Organization of the United Nations, more than 3.1 billion people around the world depend on fish for almost 20% of their animal protein, and fishing helps sustain the economies of dozens of countries. One of the major impacts on marine biodiversity is overfishing (approximately 90% of the world's fish stocks are overfished or fully fished). Not only does it lead to a disbalance in the regenerative power of species, as they cannot keep up their breeding pace with the pace and magnitude at which they are caught, it also is destructive for their habitat. Moreover, more than 15% of global fishing catch is illegal, unreported, and unregulated.[216]

Global Fishing Watch[217] produces maps of fishing effort and presence at very high resolution and publishes them as open data. Vessel tracking systems, satellite technology, cloud computing and machine learning are combined to build an accurate picture of fishing activities. This way, the global commercial fishing fleet can be monitored. The platform is helping enable scientific research, advocate for better policies to support marine protection, tackle overfishing and improve the way fishing is managed.

In collaboration with Global Fishing Watch, the government of Indonesia gained insights into Indonesian vessels that were fishing for longer than the three months allowed or in regions where they did not have a license to operate. Research by Cabral et al. (2018) shows that due to Indonesian policy to tackle illegal, unreported, and unregulated fishing, foreign fishing in its waters has dropped by more than 90% and total fishing by 25%. Their projections suggest that this could allow Indonesia's fishery economy to flourish as fish stocks regenerate.[218]

Open data can help governments make better policies to protect biodiversity.

Future outlook

Open data on biodiversity at land and sea can play a role in further developments and research that will help policymakers be more informed about biodiversity issues. In turn, better policies can be created for improving biodiversity. Biodiversity is essential to the quality of our forests and oceans. Sustaining the quality of our oceans and forests also aids in the uptake of carbon dioxide emissions, which is crucial to the climate. In the next section, we further discuss the role of open data in firstly becoming more aware of CO2 emissions and secondly to help in decreasing our carbon footprint.

214 Exchange rates as of 26 January 2020
215 Liang et al. (2016) available at: https://science.sciencemag.org/content/354/6309/aaf8957
216 FAO Report on The State of World Fisheries and Aquaculture (2016) available at: http://www.fao.org/3/a-i5555e.pdf
217 Global Fishing Watch (n.d.) available at: https://globalfishingwatch.org/
218 Cabral et al. (2018) available at: https://www.nature.com/articles/s41559-018-0499-1

5.4.5 Creating awareness about air pollution and reducing CO2 emissions

The first step in taking action to counteract global warming and reducing CO2 emissions is to make more people aware of the importance. A study in 2019 by YouGov and the University of Cambridge shows that in some large European countries, 5%-10% of people are denying climate change or think that human activity is not responsible at all (UK 5%, Denmark 5%, France 7%, Germany 7%, Norway 10%).[219]

Open data-based applications can help to create more awareness about the environmental impact we all have. In addition to merely becoming more aware of e.g. carbon dioxide emissions, applications that use open data can also encourage us to reduce our own carbon footprint. Below are several open data initiatives that:

- Make us more aware of locations with high amounts of CO2 emissions
- Make us more aware of the air quality in different locations
- Help us cope with local environmental policies such as low emission zones

A Finnish website **Päästöt.fi**[220] shows industrial emissions in all EU countries and in Iceland, Norway, Serbia and Switzerland. The website shows the emissions level on a map, based on the European Pollutant Release and Transfer Register (E-PRTR) database. The service helps to relate the orders of magnitude of emissions. It also provides citizens with the possibility to compare regions when making the decision to work or live in a region.

Plume labs[221] is an application focused on air quality. It tracks the hourly pollution levels in sixty cities in the world, including ten cities in France, four cities in Belgium and four cities in the UK. The air quality is shown using an index number, whereby above 150 is considered "critical", while anything above 100 is considered "harmful". The start-up uses the data made public by different agencies engaged in a policy of open data, for example Airparif in Paris.

The **Umweltzone app**[222] allows users to find out about the location of a low emission zone on their Android phone. Access for cars is restricted in these zones, depending on the cars' emission profile. The application uses the location coordinates of low emission zones to render them on a map. A couple of cities in Germany provide the data in a suitable, machine readable format - as open data. About half of the data is derived from OpenStreetMap since certain cities do not publish their data.

219 YouGov-Cambridge Globalism Project (2019) available at: https://yougov.co.uk/topics/science/articles-reports/2019/09/15/international-poll-most-expect-feel-impact-climate
220 Päästöt (n.d.) available at: http://paastot.fi/hiilidioksidi/
221 Plume labs (n.d.) available at: https://plumelabs.com/en/
222 Umweltzone (n.d.) available at: https://play.google.com/store/apps/details?id=de.avpptr.umweltzone&hl=en_US

As mentioned before, it is beneficial for us to be more aware of pollution levels in different areas to possibly adapt our behaviour. Open data-based apps such as CarbonCulture and Open Food Facts further aim to help people to reduce their carbon footprint and other factors that aid in a sustainable lifestyle. This more sustainable lifestyle, including reduced carbon dioxide emissions, in turn, has a positive economic effect on the environment and helps people to reduce costs on energy and water consumption.

Open Food Facts[223] is a Wikipedia-like tool for food products. The app uses EFSA data to show the public what is in the food they eat daily, covering 650 000 products from around the world. By integrating the OpenFoodTox open data on chemical hazards, it also helps identify the amounts and types of additives in food. In addition, by tracking the ingredients of consumer products with Open Food Facts, you can filter products with unsustainable ingredients such as palm oil.

CarbonCulture[224] collects national and local data on carbon emissions, energy consumption, water usage and waste disposal in Greater London. Additional data is collected from the Department for Environment, Food & Rural Affairs (DEFRA) and the Department of Energy & Climate Change (DECC). CarbonCulture aims to help people reduce carbon emissions, water usage, waste, energy consumption and save money.

[223] Open Food Facts (n.d.) available at: https://world.openfoodfacts.org/discover
[224] CarbonCulture (n.d.) available at: https://platform.carbonculture.net/landing/

Improving language services with open data

Open data helps to increase machine translation quality and reduces costs

Reduced translation cost enables especially SMEs to reach a larger audience

Open data enables equal access to language technology and services

The above estimations are potential numbers for EU28(+) or EU27(+) depending on availability of data. For details on calculations and assumptions see corresponding sections.

EUROPEAN DATA PORTAL

5.5 Language services

Over the last ten years, the global translation services market has doubled in size, reaching $46.9 billion (€42.3 billion[225]) in 2019 and estimated to reach $56 billion (€50.8 billion[226]) by 2021[227]. Europe accounts for around half of the global translation services market[228], with 25 official languages in the EU28+.[229] In general, it is important that people can use their own language when accessing content. For some people it is just not possible to consume content in another language. For others, interpretation of another language might be difficult and can pose serious consequences in case of misinterpretation.

Nowadays, translation does not only apply to traditional, printed media such as documents and books. An increasing amount of content is shared via digital channels. These can refer to the same documents and books that are now digitally available, but also to videos, mobile applications, games or e-learning courses, to name a few. In addition, consider the rise of speech recognition services, chatbots and virtual assistants. As the world is growing via digital networks and we communicate more digitally, translation services will diversify. These changes will also affect demands from public services, making translation services and language data more important in the EU.

In this section, first an explanation on how open data aids development of translation services will be given. In addition, some examples of benefits of improved translation services are elaborated upon, such as increased reach and increased access to content in one's native language.

How open data helps improve translation

Open data can help improve translations services and their algorithms. Algorithms are trained by input data. Using open data can substantially increase the amount of data used to train these algorithms without having to establish a budget for that. More high-quality data can increase the reliability and quality of translation by services based on it. So, as intelligent systems are enabled to learn from larger sets of high-quality data, it results in better machine translations (e.g. Google Translate, Microsoft Translator, DeepL, Tilde).

In addition to higher-quality input, open data makes machine learning for translation cheaper. It is cheaper because machine learning needs large amounts of data to train the program, which does not need to be purchased as open data is freely available. This way, open data is aiding the development of automatic neural machine translation, which is much cheaper than human translation. Human translation will not be replaced by neural machine translation, but it could be supported by it. Instead of having all translation work done by human translators (which could cost approximately €1.50-€2.20 per line of text or €0.15 per word[230] and produce approximately 2500 words a day or 300 words per hour[231]), at least a part of it can be done automatically or as a combination of automatic and human translation.

[225] Exchange rates as of 26 January 2020
[226] Exchange rates as of 26 January 2020
[227] Statista Market size of the global language services industry 2009-2021 (2019) available at: https://www.statista.com/statistics/257656/size-of-the-global-language-services-market/
[228] Statista Market size of the global language services industry 2009-2021 (2019) available at: https://www.statista.com/statistics/257656/size-of-the-global-language-services-market/
[229] EU 23 official languages + Norwegian and Icelandic (n.d.) available at: https://europarlamentti.info/en/European-parliament/working-languages/
[230] Approximate price of human translators in Europe (2018) available at: https://www.polilingua.com/blog/post/Translation-services-rates.htm
[231] CSA research (2019) available at: https://csa-research.com/More/Media/Press-Releases/ArticleID/38/Translation-Industry-Headed-for-a-"Future-Shock"-Scenario

For documents that contain longer sentences and context, e.g. reports or novels, there will be a need for post-editing by a translator. Research on the average increase in productivity on post-editing novels shows that translators can achieve a 36% increase in productivity when using neural machine translation[232]. Assuming this finding can be applied to governmental documents, it can be concluded that translation costs can be decreased by 36% (if the cost structure is adapted accordingly).

Open data can help decreasing tax spending on translation by 36%.

Decreasing barriers for content to be created

Lowering the costs of translation not only saves money, it will also enable more content to be translated, which was considered just too costly to manually translate before. Think of the time and effort it would take to translate huge databases such as the metadata describing all datasets on the European Data Portal or all the patents available at the European Patent Office (EPO). With open data-enabled machine learning, these databases are translated more efficiently.

EPO Patent Translate[233]: In cooperation with the national patent offices in EU member states, the European Patent Office in collaboration with Google used millions of official, human-translated patent documents. These were used to train the translation engine to handle technical subject-matter and the specific style and format used for patent documents. Every year, millions of documents are added to the EPO which too are used to continually improve the Patent Translate engine.

Open data enabled machine translation decreases barriers for content to be translated.

Increasing content available in native language

There is thus a large amount of content that could potentially be made accessible for everyone in their own language with the help of open data. Someone speaking a language that relatively few people speak, e.g. Latvian, has access to far less content than there is available in the world or even only in English.

For example, a search for "flu remedies" will result in approximately 321 000 results in Latvian, as opposed to 229 000 000 results in English (i.e., 713 times as many results in English).[234] These results might be similar for videos, news articles and academic research. When all this information can be instantly translated to Latvian, this would mean an increase of at least 713 times as much content as before. This could also be reflected in news and other media. Someone who is only exposed to news in their native language, might be biased in his/her perception of what is going on in the world. So, to get a more complete picture, it would be beneficial to have access to news in other languages and other point of views.

[232] Toral, Wieling, and Way (2018) available at: https://www.researchgate.net/publication/325152789_Post-editing_Effort_of_a_Novel_With_Statistical_and_Neural_Machine_Translation
[233] EPO Patent Translate (n.d.) available at: https://www.epo.org/searching-for-patents/helpful-resources/patent-translate.html#tab1
[234] Google Search Engine November 2019

GDELT Project[235]: Even the largest teams of human translators cannot read and translate every word published by the world's news media each day. The GDELT Translingual platform is a real-time streaming news machine translation deployment: all global news that GDELT monitors in 65 languages, representing 98.4% of its daily non-English monitoring volume, is translated in real-time into English and processed.

Another example relates to creating a larger potential audience to reach. Previously, an individual can share a YouTube tutorial in, for example Norwegian, mostly to other people speaking Norwegian. Now, with the help of real-time translation, the video can be shared with a large group of people not speaking Norwegian with instant subtitles in their native language - making it easier to reach a larger audience.

Or consider the example of enabling non-native speakers to interact with local governmental bodies or the EU. Currently about half of the people applying for a study abroad in the EU could not find relevant information for applying due to language issues.[236] Governmental bodies can make procedures for studying abroad, but also procedures for taxes or permits available in all languages and localised based on context - making it easier for non-native speakers in a country such as exchange students, expats, or immigrants to understand and participate. These are thus ways in which open data enhanced neural machine translation can help remove language barriers.

Open data helps to increase the amount of content available in someone's native language.

Future outlook

The previously mentioned prospected growth in the translation services market can be partly attributed to a greater variety of content to be translated and more widespread use of neural machine translations (NMT). These have been developed so well that *"by 2022 most business translations will be carried out by NMT engines, which will be largely near free-of-charge, with human post-editing"* says CEO of one of the largest global translation companies.[237] Translation companies could thus benefit from combining the work of human translators and machine translation for the growing demand in translation services. As the cost of initial translations by neural machine translations will become near zero, barriers for content to be translated fade. This might open a target group, which otherwise would perhaps not have its contents translated because of the costs or the time it takes to have everything translated by humans.

To conclude, improved machine translations can, in combination with a human final check, be applied to a larger amount of content at a fraction of the cost and time of human-only translations. In other words, saving money as opposed to current practices, while at the same time making more content accessible. Moreover, open data does not only allow current players in the market to benefit but can also enable smaller companies to develop their solutions, thus playing a potential key role in further development of new solutions for overcoming language barriers. Supporting SMEs for a thriving economy and fostering a Europe that is growing together without losing its variety in languages and cultures is one of the goals of the European Union. Open data actively supports those goals in enabling better transfer of knowledge and more equal access to language technology.

235 GDELT Project (n.d.) available at: https://www.gdeltproject.org/
236 Capgemini research (2019)
237 Ofer Shoshan quote on neural machine learning (2019) available at: https://www.verdict.co.uk/ai-translation-nmt/

6. SAVING COSTS DUE TO OPEN DATA

6.1 Intro and methodology

In addition to adding value through lives saved, time saved, environmental benefits, and increasing knowledge transfer, there are also costs saved with the help of open data. These can refer to cost savings for organisations that simply open up their own data, cost savings thanks to acquiring others' open data for free or at marginal cost, and cost savings through efficiencies enabled by open data re-use.

Cost saved by publishing open data

There are benefits and cost savings for organisations that make their data openly available according to a literature research by Welle Donker (2018). First, data quality will improve thanks to the feedback of the re-users, who will need to assess its quality and address its issues before they can use it in their own applications. In Denmark, for example, more feedback was received after making address data owned by the Danish government publicly available as part of crowdsourced map service OpenStreetMap[238].

Secondly, administrative and compliance costs related to data sharing obligations (e.g. freedom of information requests to government bodies) as well as the costs related to managing restrictive data sharing terms (e.g. bilateral, possibly non-standard contractual agreement with re-users) can significantly decrease. The research by Welle Donker et al. (2018)[239] found that in the Netherlands, government organisations providing open data indicated that they had received fewer requests under the Public Information Access Act in respect to organisations that do not, which would indicate lower administrative costs in addressing those requests. Also, the implementation of open data policy in 2011 by the Dutch Education Service (DUO) led to a 60% reduction of formal requests, saving DUO around 4.5 FTE in two years[240].

Cost saved by acquiring open data for free or at marginal cost

On the other side, significant costs can be saved by re-using open data - available for free or at minimum cost from public administrations - that otherwise would need to be purchased at market price. The Dutch National Institute for Cultural Heritage, for example, expects that municipalities will save around seven minutes per application for spatial planning because the necessary data is openly available[241]. Another aspect of the value of open data refers to the costs that are directly saved by opening up specific datasets to be accessed free of charge or at marginal costs. There used to be considerable revenues generated by PSI suppliers as users were charged for access to the datasets. Costs can be saved when re-using open data instead of paying for the data that is commercially published. In the Netherlands for example, users were charged approximately €63-€68 million for using the datasets from PSI providers such as the Dutch Chamber of Commerce, the Cadastre, the CBS (the Dutch national statistics agency) in 2009-2010[242]. As these datasets are now freely available, users can save themselves these costs. Moreover, more people and organisations can start to use this data as the financial barrier is now removed.

238 Welle Donker (2018) available at: https://www.researchgate.net/publication/328501181_Funding_Open_Data
239 Welle Donker (2018) available at: https://www.researchgate.net/publication/328501181_Funding_Open_Data
240 Kronenburg et al. (2012) available at: https://kennisopenbaarbestuur.nl/media/23577/de-waarde-van-open-data.pdf
241 Kronenburg et al. (2012) available at: https://kennisopenbaarbestuur.nl/media/23577/de-waarde-van-open-data.pdf
242 Vickery (2011) available at: http://ec.europa.eu/newsroom/document.cfm?doc_id=1093

Cost saved through efficiencies enabled by open data

In addition to the cost savings described previously, with the availability of open data, organisations will be able to operate more efficiently and effectively as seen in the previous chapters. Time and money are the most important factors when looking at the economic impact. In this report - when it adds value - the value of time is expressed as money, knowing however that time is not an equivalent to it. In general, the report refers to costs saved in euros. If secondary research is used, the exchange rates of January 2020 are used to calculate costs in euros.

Each of the topics discussed previously, i.e., lives saved, time saved, environmental benefits, and access to content through translation, can have associated indirect cost savings. Although these benefits are not always easy to quantify into monetary benefits directly, some examples are provided below.

Saving costs with open data

€3.7 – €5.2 billion healthcare costs savings

due to better allocation of resources to combat malaria globally*

€312 - €400 thousand healthcare cost savings

because first responders can administer CPR before the arrival of emergency services

* Potential lives saved due to enhanced malaria treatment are global numbers.

The above estimations are potential numbers for EU28(+) or EU27(+) depending on availability of data. For details on calculations and assumptions see corresponding sections.

EUROPEAN DATA PORTAL

6.2 Saving costs in healthcare and emergency services

6.2.1 Saving costs in malaria treatment

According to UNICEF, the costs of malaria medication is approximately $2.00-$2.50 (€1.81 – €2.27[243]) per adult treatment dose[244]. Research by Ezenduka et al. (2017)[245] found that in Nigeria the treatment of uncomplicated malaria by a public health facility costs approximately $31.50 (~€28.57[246]) per case (82.5% is personnel costs and 6.6% malaria medication). Considering the potential decrease of 131.40 - 186.15 million cases of malaria with the help of open data-based tools, this would lead to a potential cost reduction of at least $262.8 – $372.3 million[247] (€238.4 – €337.7 million[248]) on malaria medication alone. If the total expenses by health facilities is considered – assuming the estimation in Nigeria – globally a potential of $4.1 – $5.8 billion (€3.7 – €5.3 billion[249]) could be saved each year.

> Open data can help decrease malaria infections resulting in potential annual healthcare costs savings of €3.7 – €5.3 billion globally.

6.2.2 Saving costs due to bystander CPR

The life-saving medical procedure cardiopulmonary resuscitation (CPR) by bystanders is crucial to the survival of people having an out-of-hospital cardiac arrest (OHCA). Research shows that bystander CPR was positively associated with long-term survival.[250] As described in the chapter on lives saved, open data-based apps that guide bystanders to support a person with an out-of-hospital cardiac arrest can help increase the number of people receiving bystander CPR from 140 to 235 thousand across Europe, i.e., 95 thousand more people.

Research by Riddersholm et al. (2017)[251] found that the amount of days spent in a hospital after out-of-hospital cardiac arrest is lower for people who receive CPR by bystanders before an ambulance arrives. Their results show an average amount of days spent in the hospital of 20 days in case of no bystander intervention, 16 days if bystander CPR was administered, and 13 days if people were defibrillated by bystanders. Another research by Tan et al. (2017)[252] shows that the average patient-related care costs per day in the intensive care unit is €1040 in Germany, €1243 in the Netherlands, and €1333 in Italy.

> Open data can help more people get bystander CPR resulting in potential annual healthcare costs savings of €312 - €400 thousand in EU28+.

[243] Exchange rates as of 26 January 2020
[244] UNICEF Malaria Fact Sheet (n.d.) available at: https://www.unicef.org/media/media_20475.html
[245] Ezenduka et al. (2017) available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5691839/
[246] Exchange rates as of 26 January 2020
[247] Considering $2.00 cost per adult treatment dose
[248] Exchange rates as of 26 January 2020
[249] Exchange rates as of 26 January 2020
[250] Bürger et al. (2018). Study of 10853 out of hospital cardiac arrests in Germany 2010-2016. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6156551/
[251] Riddersholm et al. (2017) available at: https://www.resuscitationjournal.com/article/S0300-9572(17)30295-2/fulltext
[252] Tan et al. (2017) available at: https://zfin.org/ZDB-PUB-170127-2

Saving costs with open data

€739.8 million labour costs saved

due to reduction of waiting time because of real-time notifications in case of delayed trains

€13.7 – €20 billion labour costs saved

due to reducing time stuck in traffic with real-time traffic navigation

The above estimations are potential numbers for EU28(+) or EU27(+) depending on availability of data. For details on calculations and assumptions see corresponding sections.

EUROPEAN DATA PORTAL

6.3 Saving costs in public transport and traffic

6.3.1 Saving costs due to less time spent in public transport

Open data applications can notify people of the accurate departure times of public transport. In the time saved chapter, it is estimated that providing real-time public transport information with open data-based apps can help save 27 million hours annually for all European train travellers. The lives of travellers would be substantially impacted. To make a financial estimate of the value of these 27 million hours saved, we can use the average hourly labour cost in 2018 of €27.4 in the EU28[253]. This leads to an approximate value of €739.8 million total. The time saved could be used either productively or as relaxing time, i.e., sit a little longer on the couch before you leave your home, or relax a bit and detach from your tasks after a working day.

> Open data can potentially save €739.8 million due to time saved in public transport.

6.3.2 Saving costs due to less time spent in traffic

Research on the 200 most congested cities worldwide shows that drivers spent on average 129.5 hours stuck in traffic in European cities. In German cities, for example, road users lost on average 120 hours, costing approximately €1050 per driver[254]. In total, it is estimated that Europe's traffic hotspots could cost drivers €217.14 billion[255] by 2025 due to time wasted in traffic.[256]

> Open real-time traffic data can potentially save €13.7-€20 billion due to reduced waiting times in traffic.

Open data can help save time in traffic by providing real-time traffic data, based on which road users can adapt their route in case of congestion. It is most likely that real-time car navigation apps such as Google Maps or Waze are already realising this value, using open data as one of the sources to their algorithms in those countries where it is available. In the time saved chapter, it is estimated that this results in approximately 500 – 730 million hours potentially saved each year by European drivers commuting to work in urban areas. Based on the average hourly labour cost in 2018 of €27.4 in the EU28[257], this would lead to an approximate potential cost saving of €13.7 - €20 billion in labour costs.

253 Eurostat Labour Costs (2019) available at: https://ec.europa.eu/eurostat/statistics-explained/index.php/Hourly_labour_costs
254 INTRIX Global Traffic Scorecard (2018) available at: http://inrix.com/scorecard/
255 Exchange rates as of 26 January 2020
256 INTRIX Europe's Traffic Hotspots: Measuring the Impact of Congestion in Europe (2016) available at: http://www2.inrix.com/traffic-hotspots-research-2016
257 Eurostat Labour Costs (2019) available at: https://ec.europa.eu/eurostat/statistics-explained/index.php/Hourly_labour_costs

Saving costs with open data

€5.1 billion costs saved on energy bills

due to reduced household energy consumption

€79.6 billion costs saved in electricity bills

because of efficient use of solar panels on rooftops

The above estimations are potential numbers for EU28(+) or EU27(+) depending on availability of data. For details on calculations and assumptions see corresponding sections.

EUROPEAN DATA PORTAL

6.4 Saving costs due to environmental benefits

6.4.1 Saving costs by reduced energy consumption

In the EU28 countries, energy consumption by residential households account for approximately 288.3 million tonnes of oil equivalent (Mtoe), in 2017[258]. Open data-based applications and services such as Opower[259] (see chapter 5.4.1) are assumed to serve as a tool to help European households reduce their energy consumption by approximately 2%[260]. The potential energy saved by households with the help of open data would be approximately 5.8 million tonnes of oil equivalent each year. Of the total energy consumption of households, on average 24.1% is electricity and 36% is gas[261]. If the potential 5.8 Mtoe saved with the help of open data reflects the same proportions, this would mean that 24.1% (1.4 Mtoe or 16282 GWh) is electricity and 36% (2.1 Mtoe or 24423 GWh) is gas[262].

Based on the average costs in the EU28 countries of €0.2159[263] per kWh electricity and of €0.0632[264] per kWh gas, the potential cost saved by households for electricity and gas alone could be approximately €5.1 billion annually.

> Open data can potentially help households reduce their energy bills by €5.1 billion in EU28.

6.4.2 Saving costs by increased solar energy

In the environmental benefits chapter, it is shown that open data can be used to provide information on the optimal use of solar panels on rooftops. In the EU28 countries, there is a total rooftop area available for solar panels of 7935 km2. Bodis et al. (2019)[265] studied the extent to which this rooftop area can be used to significantly increase the production of solar energy across Europe. They estimate levelized cost of producing electricity by solar panels[266] in each country to be between €0.0619 and €0.3215 per kWh. The research found that the potential for rooftop solar systems that produce electricity at a lower cost than the retail electricity prices in each country, i.e., the economic potential, is approximately 467 TWh in the EU28.

> Open data has the potential to help save up to €79.6 billion annually in electricity bills by increasing solar energy production.

258 Eurostat Energy Consumption (2019) available at: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Energy_consumption_in_households#Energy_products_used_in_the_residential_sector
259 Energy.gov (2013) available at: https://www.energy.gov/data/articles/who-uses-open-data
260 Allcott and Rogers (2014) available at: https://www.povertyactionlab.org/evaluation/opower-evaluating-impact-home-energy-reports-energy-conservation-united-states
261 Eurostat Final energy consumption in the residential sector by fuel, EU-28 (2017) available at: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Energy_consumption_in_households#Energy_consumption_in_households_by_type_of_end-use
262 Unit converter Mtoe to GWh (n.d.) available at: https://www.iea.org/statistics/resources/unitconverter/
263 Eurostat Gas prices for household consumers (2019) available at: https://ec.europa.eu/eurostat/statistics-explained/index.php/Natural_gas_price_statistics
264 Eurostat Electricity prices for household consumers (2019) available at: https://ec.europa.eu/eurostat/statistics-explained/index.php/Electricity_price_statistics
265 Bodis et al. (2019) available at: https://www.sciencedirect.com/science/article/pii/S1364032119305179
266 Levelised cost of producing electricity by solar panels is calculated using capital investment, operation and maintenance costs, electricity generation, discount rate, and investment period considered (for more detail see Bodis et al. (2019) available at: https://www.sciencedirect.com/science/article/pii/S1364032119305179

If we assume that the difference between the retail price of electricity per kWh and the cost of producing the same by solar panels is at maximum the difference between the average retail electricity price in each country and €0.0619 (the lowest levelised cost of producing electricity as defined by Bodis et al. (2019)), and multiply that by each country's economic potential, a potential cost saving of up to €79.6 billion each year could be realised in the EU28 countries combined. Open data could play a role in increasing the solar energy production to the economic potential of 467 TWh defined by Bodis et al. (2019). These benefits might also be achieved without open data. However, open data provides the necessary information free of charge, enabling services (e.g. Sunenergia[267]) to make use of this information and thereby making it easier for households, businesses, and governments to use their buildings' rooftops for solar panels.

[267] Sunenergia (n.d.) available at: https://sunenergia.com/en/

Saving costs with open data

€0.25 - €1.48 billion costs saved

due to open data in the public sector

€1.1 billion translation costs saved

due to the help of machine translation in translating all datasets on the European Data Portal

The above estimations are potential numbers for EU28(+) or EU27(+) depending on availability of data. For details on calculations and assumptions see corresponding sections.

EUROPEAN DATA PORTAL

6.5 Saving costs due to increased knowledge transfer

6.5.1 Saving costs in public sector

Open data can help government bodies to operate more efficiently and effectively. Time, and in turn, costs are saved, which can be reallocated to other processes. In Denmark, for example, research shows the potential cost saving for the public sector due to opening up data. The Danish Government set out a public sector-digitisation strategy to provide what they call "basic data" - the core information government authorities use in their day-to-day operations - from all layers of government (ministries, municipalities and regions). Between 2012 and 2020, certain categories of basic data were gradually made freely available to all public authorities, private companies and individuals. In a report from 2012, the foreseen government savings are considered to be around 35 million EUR per annum, starting from 2020[268].

A more recent Danish study (2017)[269] about the impact of open geospatial data refers to the effect of open data on reducing time-consuming processes. In this study, it is estimated that public authorities save 22 million DKK (€2.94 million) thanks to the release of geospatial data. These figures can be extended to the whole Danish Open Data market, "as spatial information makes up about one half of all PSI according to various estimates" (see e.g. PIRA (2000)[270]; MEPSIR (2006)[271]; Vickery (2011)[272]). For public sector saving, this would account to approximately €5.9 million annually.

Based on this estimate and the numbers of the 2017 open geospatial data the government savings are expected to be in a range from €5.9 - €35 million each year. The expected cost savings for the Danish government in 2020 can be expressed as a percentage of the total public sector expenditure using the latest (2018) GDP and government expenditure data. In 2018, the Danish government expenditures accounted for 51.5% of the approximate €320 billion GDP, i.e., €165 billion. The expected cost saving ranging from €5.9 to €35 million, is then approximately 0.004% to 0.021% of government expenditure. Under the assumption that the potential cost saving in Denmark (as a percentage of government expenditure) can be applied to Europe, approximately €0.25 - €14.8 billion can be saved annually in the EU27+ public sector (based on 2018 government expenditures as % of GDP).

> Open data can create a potential cost saving of €0.25 - €14.8 billion in the public sector across Europe.

6.5.2 Saving costs in public sector due to improved machine translation

Open data can help make services of the public sector more efficient, for example by aiding in translation. In the EU28+, there are 25 official languages.[273] It is important that people can use their own language when dealing with the EU and when accessing content in general. As mentioned in the section on language and translation, open data is aiding the development of automatic neural machine translation, which is much cheaper than human translation. Machine translation is not intended to completely replace human translation, but it can support human translators. Instead of having all translation work done by human translators (which

[268] The Danish Government (2012) available at: https://en.digst.dk/media/18773/good-basic-data-for-everyone-a-driver-for-growth-and-efficiency.pdf
[269] Danish Agency for Data Supply and Efficiency (2017) available at: http://sdfe.dk/media/2917052/20170317-the-impact-of-the-open-geographical-data-management-summary-version-13-pwc-qrvkvdr.pdf
[270] PIRA (2000) available at: http://ec.europa.eu/newsroom/document.cfm?doc_id=1195
[271] MEPSIR (2006) available at: http://ec.europa.eu/newsroom/document.cfm?doc_id=1197
[272] Vickery (2011) available at: http://ec.europa.eu/newsroom/document.cfm?doc_id=1093
[273] EU 23 official languages + Norwegian and Icelandic (n.d.) available at https://europarlamentti.info/en/European-parliament/working-languages/

could cost approximately €1.50-€2.20 per line of text or €0.15 per word[274] and produce approximately 2500 words a day or 300 words per hour[275]), at least a part of it can be done automatically or as a combination of automatic and human translation.

> Open data can aid development of machine translations resulting in a potential cost saving of €1.1 billion in translations on the European Data Portal.

Consider the following example of translating the metadata of the datasets available on the European Data Portal. Currently, around 1 million datasets are available through the portal, each accompanied by metadata that contains on average around 300 words each. The metadata is available in all the 25 official EU languages. It would take a human translator about one hour to translate the metadata of one dataset from the original language to each of the other 24 languages, and thus 24 million hours to translate the whole European Data Portal. These hours combined are 3 million working days which would mean it takes a single translator 120 000 years to translate the metadata for all datasets. Based on the cost of a human translator, the costs of translating all metadata of the whole European Data Portal could be 300 words in 24 translations for 1 million datasets times €0.15, resulting in approximately €1.1 billion. Thus, translating the European Data Portal would simply be impossible in terms of time, and unacceptable in terms of cost, if not thanks to the employment of neural machine translation.[276]

[274] Approximate price of human translators in Europe (2018) available at: https://www.polilingua.com/blog/post/Translation-services-rates.htm
[275] CSA research (2019) available at: https://csa-research.com/More/Media/Press-Releases/ArticleID/38/Translation-Industry-Headed-for-a-"Future-Shock"-Scenario
[276] Note: not including development, training, and operational costs of the neural machine translation tool.

7. OPEN DATA IN ORGANISATIONS

To understand the economic impact of open data in organisations, companies across Europe were surveyed, researched and interviewed.

In order to achieve statistical significance, organisations were initially selected based on a key controlling for geographical and sectoral spread using Thomson Reuters' Eikon database. 2500 companies were privately invited to partake in the survey. Afterwards, a link to the online survey was also made available for any other company for a period of three months until the 29th of August 2019. Three of the actively invited companies took part in the survey. 103 volunteer companies participated via the open link.

Due to the statistically insignificant size of the sample that was obtained from the closed survey, the 106 responses were analysed as a whole. Therefore, the opportunity to mitigate geographical or sectoral bias was not available. It was taken into account that the sample of participants was no longer representative of a European overview of organisations. In addition, organisations that took part in the survey voluntarily are most likely more mature and/or exposed to open data, which influenced their answers. The resulting bias was taken into consideration. In any case, it was observed that significant insight into the perception around the economic impact of open data in organisations can be inferred from the survey participants' responses. The key findings are summarised below.

Finally, 28 of the 103 respondents were selected to serve as case studies. Their open data stories are presented in 28 brief case studies which can be found in the appendices. Of these, three previews are presented below, followed by the survey findings.

7.1 Open data stories from European organisations

France - Lefebvre Sarrut

Lefebvre Sarrut is a legal publisher that operates in several European countries. Their main product is a legal doctrine where they comment on the law and jurisprudence (primary sources by lawyers and academics). In half of the countries they operate in, this legal data is distributed through open data. There are also other very specific datasets that they leverage from open data, like collective agreements in labour law (i.e. France) or the tax doctrine from the tax administration (i.e. Belgium). Lefebvre Sarrut is also interested in specific economic datasets that are now available through open data, such as data describing the company activities and financial indicators, which are especially valuable when combined with legal data. When open data is not available, they need to acquire the data themselves - sometimes buying it from public administration - and prepare it.

In some countries, Lefebvre Sarrut helps administrations increase the quantity of open data, such as their recent collaboration with the French administration on jurisprudence. In addition, they are highly involved in open source and open data communities. They have been involved in an association called Open Law (le droit ouvert) since 2015 and have been taking part in, or running, some open innovation programmes where public administrations (e.g. Supreme Court, Legifrance, etc.) and private companies collaborate to develop open data. In association with Open Law, they are discussing how to rationalise their efforts and share work. They hope that these discussions will positively influence policymakers to make it easier for stakeholders to act together and share their work. Lefebvre Sarrut expects a significant increase along with the full implementation of open data in the future. "*L'avenir ne sera pas ce qui va arriver mais ce que nous allons faire (the future is not what will happen but what we will do)*" - Henri Bergson.

Poland - City of Warsaw

The City of Warsaw has been using open data since 2014 with their Apps4Warsaw project - an initiative in partnership with the City Hall, Warsaw University of Technology and a commercial ICT company. The goal of the project was to develop applications based on data published on the Warsaw Open Data Portal, which was launched as a first step of the project. The City of Warsaw shares more than 200 data archives and collections through the city's open data portal, which serves about 1 million database queries per day. Several dozens of applications were built by start-ups, companies, and developers. Some of those projects achieved commercial success and are still being developed. Yet, measuring the social and economic impact of open data is still a challenge in Warsaw.

The main challenge that the City of Warsaw is facing regarding open data is the siloed solutions that make integrating data more complicated. This is more an organisational than a technical issue, caused by how internal data owners (departments and municipal companies) do not always fully understand the necessity of opening data and the value of doing it in a standardised way and publishing on suitable venues such as open data portals. Even if this is done, publishing open data is not usually a priority in their agendas. Nevertheless, the City of Warsaw believes that the usage of open data will increase. They want to give open access to their city data and to datasets provided by external entities (using API standards). Moreover, they want to give information about ICT devices and technologies they use, about conditions of using the city's ICT infrastructure and communication standards. They hope that these efforts will bring them closer to realising the full potential of the local data market. The City of Warsaw believes that creating a local data market will yield wider re-use of open data. Re-use is further enabled by a series of training courses on open data and smart cities that they organised in 2019 and will be continued in 2020.

The Netherlands – Open State Foundation

Open State Foundation is a non-governmental organisation (NGO) that promotes digital transparency of the public sector. They do this by advocating that governments release more open data and promote its re-use amongst key stakeholder groups, such as journalists, companies, other government agencies and NGOs. Open State Foundation primarily uses open data to document election results and government decision-making, public spending and performance in general. They also promote the release of government-owned data such as public transport data and the company register that are not freely available yet in the Netherlands.

Moreover, Open State Foundation also uses open data to build platforms, such as *openspending.nl*[277] to compare spending of municipalities, *waarismijnstemlokaal.nl*[278] for finding a polling station during elections, or zoek.*openraadsinformatie.nl*[279] for searching minutes from municipalities and provinces meetings. They also advise government on how to release information or improve the data quality, build platforms with the data and promote the re-use of this data through events, such as hackathons and app challenges.

[277] Open spending (n.d.) available at: https://openspending.nl/
[278] Waar is mijn stemlokaal? (n.d.) available at: https://waarismijnstemlokaal.nl/
[279] Open raadsinformatie (n.d.) available at: https://zoek.openraadsinformatie.nl/

Open data in organisations

103 organisations participated in the survey

21 European countries are represented

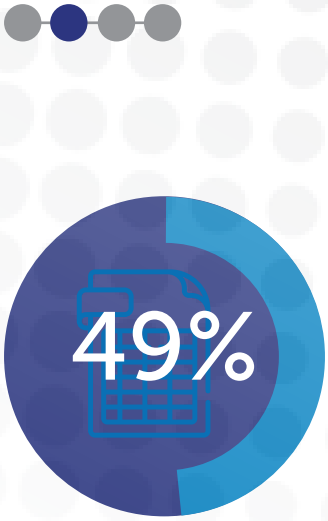
Including non-EU countries United Kingdom, Switzerland, and Serbia

Most organisations (31%) are active in the information and communication sector

Description of survey participants of EDP research (2019)

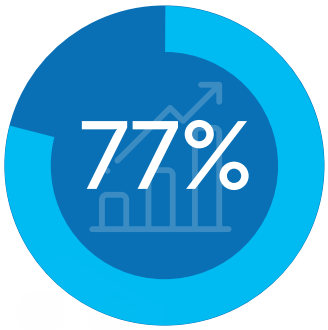
EUROPEAN DATA PORTAL

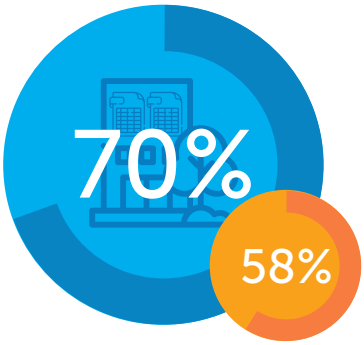
Open data in organisations



49% of data used by organisations is open data

77% of organisations plan to use more open data





70% of organisations create data internally

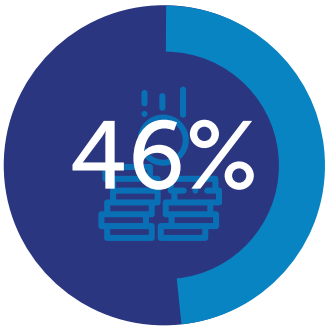
of which 58% publish some of it as open data

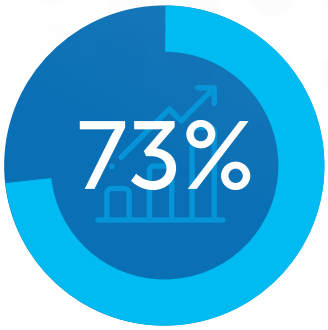
Survey results from EDP research (2019)

EUROPEAN DATA PORTAL

Open data in organisations

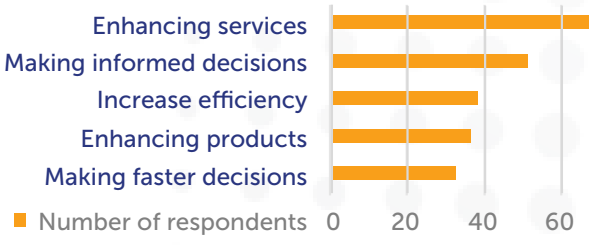
46% of organisations' revenues are impacted by open data

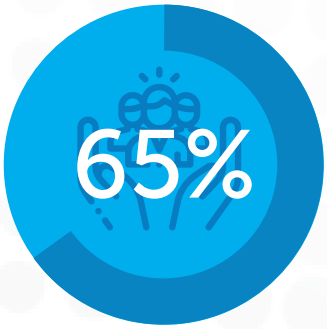




73% of organisations expect the impact of open data to increase in the next 5 years

65% of organisations use open data to enhance services

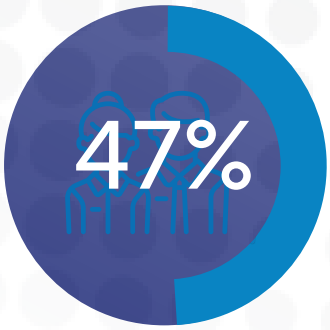




Survey results from EDP research (2019)

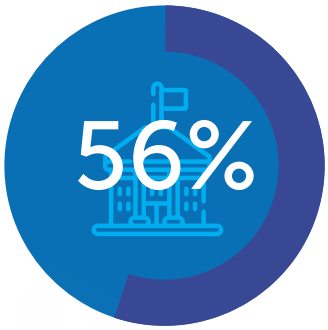
EUROPEAN DATA PORTAL

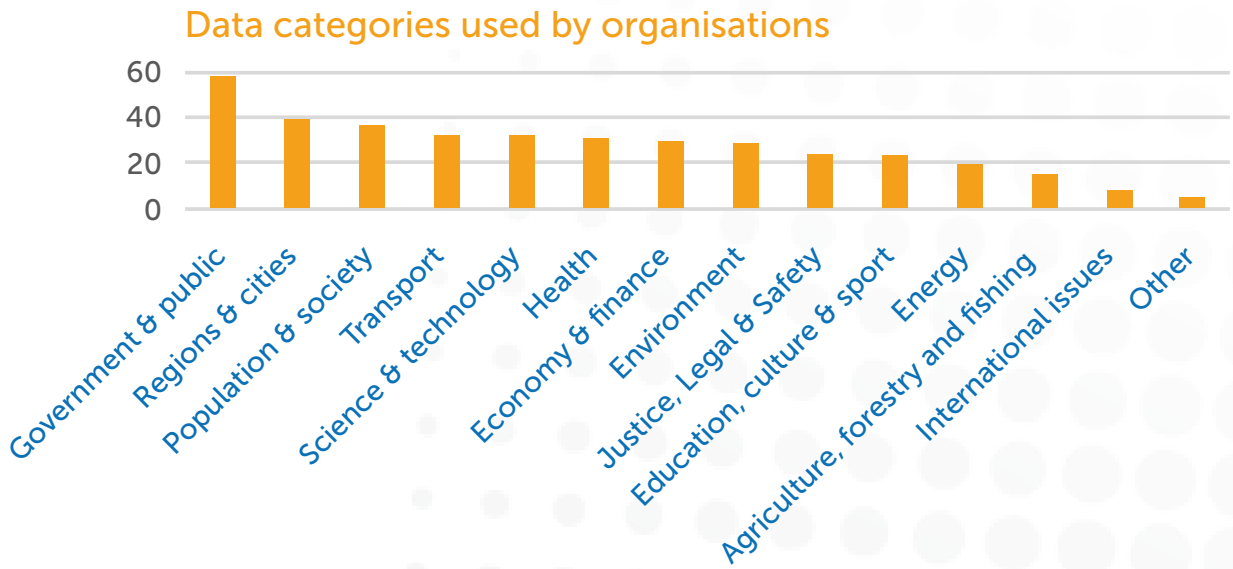
Open data in organisations



47% of employees work directly or indirectly with open data

Open data describing public sector activities are re-used most by organisations





For more information on the data categories please visit EDP website.
Survey results from EDP research (2019)

EUROPEAN DATA PORTAL

8. DISCUSSION

8.1 Summary of the results

As we have seen in the previous chapters, open data can be beneficial in a variety of sectors, in a variety of ways and for a variety of (re-)users, like businesses and governments as well as citizens. Open data enables faster and easier access to more information, which enables better decision making. Moreover, re-using open data can improve existing products and services or be used to develop new ones. Especially for small and medium enterprises and start-ups, which normally may not be able to acquire access to data or generate similar data themselves, open data opens up a range of opportunities.

Open data market size

Findings from relevant literature show a narrow range of results on open data market size when made comparable as share of GDP. Applying the median share % (1.19%) from these studies to the current GDP suggest a current open data market size of €184.45 billion for EU27+. In the baseline scenario, the open data market size for EU27+ will reach €199.51 billion in 2025. In the optimistic scenario it is €334.20 billion in 2025. This results in a growth potential of €134.69 billion.

Open data employment

Growth is also expected in the number of people who are employed due to open data. The chapter on open data employment shows that the current number is approximately 1,09 million. Over the next years, more people are expected to be working in open data related jobs and especially in jobs indirectly created due to open data. It is forecast that in 2025 the number of direct and indirect open data employees can potentially grow to 1.97 million in an optimistic scenario. This is a growth potential of 884 thousand open data employees. When by 2025 the value created by each employee exceeds €169 thousand (current status in 2019), the open data market size could potentially be even larger than forecast.

Open data potential per sector

The sector view on the open data market aimed at understanding if and in which sectors the open data market potential can be realised. When exploring open data value creation in different sectors, differences in growth potential can be observed. Ten sectors are expected to shoulder most of the growth in order to further leverage the open data impact potential.

- High impact sectors: Public administration; Professional, scientific and technical activities; Information and communication and ICT; Transportation and storage
- High potential sectors: Agriculture; Financial services and insurance; Health; Education; Wholesale retail and trade; Real estate activities

There are significant differences in value creation per employee in different sectors.

Efficiency gains due to open data

The value created by open data goes beyond solely financial benefits, as we have seen in the chapter on efficiency gains. The value of open data can be considered as benefits that come directly from using it, but also indirectly because of new insights, products and services, and improved processes that are created through re-use. Not only does open data help businesses and governments generate more revenues due to new services, or reduce costs by working more efficiently, it can also help to save lives, save time, preserve the environment, or enhance knowledge transfer through language services. Several examples of benefits that go beyond financial value have been calculated as a potential for Europe. The exploitation

of these benefits today is unlikely fully tapped. For this to change, many opportunities are available and depend on open data becoming available throughout Europe, building on and extending the success stories described in the many example provided in this.

Saving costs due to open data

Each of the topics discussed in the chapter of efficiency gains, i.e., lives saved, time saved, environmental benefits, and improved knowledge transfer through language services, can have associated cost savings. Although these benefits are not always easy to quantify into financial terms, the chapter on cost saving has provided several examples. To evaluate their economic impact, cost savings are calculated for healthcare, time saved in traffic and public transport, energy bills, and public sector.

Open data in organisations

To understand the economic impact of open data in organisations, 103 companies across Europe were researched. Their open data stories are presented in 28 brief case studies which can be found in the appendices. Three previews of open data case studies are presented in the chapter on the economic impact of open data in organisations, as well as the findings of the survey they took part in. Already, almost half of the data used by these organisations is open data. This number is expected to grow, as 77% of organisations plan to use open data even more. The main reason for using open data is to enhance services, followed by making informed decisions, and increasing efficiency. Approximately 46% of revenue is currently impacted by open data, and 73% of organisations expects the impact of open data to increase in the next 5 years. In these organisations, 47% of employees work directly or indirectly with open data. The organisations often create data internally and in 58% this generated data is published as open data.

8.2 Opportunities for reaching the full potential of open data

To understand better how to reach the potential of the open data market, the concept of opportunities for value creation was introduced and applied in chapter 2 and 3 and exemplified in chapter 4 and 5. To recap, we have a potential of €134.69 billion that 1.97 million employees could create with an average value creation per employee of €169 thousand.

We can also conclude that most of this value will be created in ten either high potential or high impact sectors. Value creation in those sectors differs, based on their labour intensity. Depending on the value creation per employee, the total impact realised will change and affect open data employment.

When looking at open data value creation as a whole, opportunities might not lie only in the size of the workforce but also in the way value is and can be created in the future. It highlights that different opportunities for growth need to be explored to enable value creation with open data. Aspects that affect value creation can be found in an adequate supply of open data and successful re-use approaches. It should also be said that open data value creation is often not quantifiable for example because it is related to social, pro bono, altruistic or charitable work. In other words, open data creates value which cannot always be linked to revenues but is indeed most valuable. The above-mentioned aspects are briefly discussed in light of the report's findings in the following chapters

8.2.1 Adequate supply of open data

The amount of open data currently available is still limited (see Open Data Barometer 2018) Moreover, there are large differences in the extent to which different organisations but also governments make their data openly available (see EDP Open Data Maturity report 2019). There is a large amount of data that is suitable to be published as open data, but is not open yet, creating a missed opportunity. At the same time, many datasets are simply not suited to be openly available in their original form, because they are confidential and/or contain personal data such as medical records and require processing such as aggregation or anonymisation before they can be distributed.

In addition, there exist barriers to re-use such as low data quality, unclear and unstandardised metadata describing it, lack of clarity about its licensing, lack of suitable infrastructure enabling discoverability and accessibility (see the European Data Portal's (EDP) report on barriers of PSI re-use in the Public sector report for more detail). In order to reach the full economic value for Europe, the quantity of open data as well as the quality need to improve. This is, however, a large-scale operation. Therefore, it makes much sense to target efforts to further increase first the impact of open data on datasets that have the highest potential for both society and the economy.

In the European Directive on open data and the re-use of public sector information, also known as the Open Data Directive (Directive (EU) 2019/1024), the European Commission is tasked to specify high-value datasets (HVDs) that organisations in the scope of the Directive will have to make available free of charge, in machine-readable format and via APIs. A thorough analysis can identify high-value datasets (HVDs) and their respective specifications to create impact. Given the many sector-specific differences, this process is likely to be more successful when including sector experts familiar with data re-use.

The specification and implementation of high-value datasets as part of the new Open Data Directive is a promising opportunity to address quality & quantity demands of open data.

8.2.2 Innovative re-use approaches

Even if the quality expectations of open data are met and the quantity of open data will significantly increase, it might not be enough. Aspects like awareness and capabilities, initiatives and collaborations in and across the private and the public sector, and legal and technology frameworks enabling the combination of open data with data that is not open, are vital to achieve open data growth.

Addressing quality & quantity demands is important, yet not enough to reach the full potential of open data.

Awareness and capabilities of re-users

Many skills needed for open data re-use are not different from those needed to process any other data. Lack or mismatch of skills, in fact, is a common barrier hindering multiple aspects of digital ways of working (See for example "Data Workers Skills Gap" in IDCs European Data Market study IDC (2017). In many cases, it is a lack of capability, which extends beyond missing skills and also includes awareness, literacy, attitude and mind-set. To increase value creation due to open data, whether in a fast or slow-growing workforce, capabilities have to fit the task. Therefore, a structural way of developing competence and adapting mindset must follow the pace of digital transformation.

Open data re-users have to be aware and capable of understanding and leveraging the potential.

Even if awareness, competence, and mind-set are present, lack of incentivisation and enablement for learning and development almost always leads to reprioritisation of learning and applying new ways of working. Insufficient development and acquisition of talent, procurement of tasks or forms of collaboration hinder value creation. Hence, apart from the need for digital skills, there is a fundamental need for organisational capabilities to manage new digital ways of working and collaborating.

Open data value creation is part of the wider challenge of skill and process transformation: a lengthy process whose change and impact are not always easy to observe and measure.

Initiatives and collaborations in and across the private and the public sector

There is a huge untapped potential for publishing open data that is generated by private businesses and in so-called public-private partnerships (PPPs). Public-private partnerships refer to a, typically long-term, collaborative agreement between two or more private sector businesses and governmental bodies. They are often aimed at providing goods or services to the public. For example, public-private partnerships have the goal of financing, building, and operating projects, such as public transportation networks. Sharing data is essential in these public-private partnerships, which can be a first step to making the data openly available. In order to reach the full potential of open data, organisations and governments alike need to be aware on the benefits of open data, learn from each other, enable and incentivise exploring the opportunities for opening up datasets, and including respective licence agreements in PPP contracts. Especially intra-sector collaboration must understand and address sector-specific challenges and strategies.

Sector-specific initiatives and collaboration in and across private and public sector foster value creation.

Going beyond open data combining it with data that is not open

A shift is needed towards looking at open data holistically as part of the greater data sharing landscape. Capabilities, initiatives, and collaborations do not need to solely focus on open data re-use. In fact, a broader data re-use approach holds more potential in scope and benefit. The combination with, for example, personal data generated from devices such as fitness monitors or connected toothbrushes, or with crowdsourced data from car navigation users, or medical data from personal health accounts can be highly beneficial.

One example is Waze[280] which combines open maps data with crowdsourced traffic data. Another is Advanced Mobile Location (AML)[281] enabled phones which automatically activate the phone's location capability when making an emergency call. The latter is estimated to potentially save 800 extra lives every year thanks to the improvement in emergency caller location solutions.[282] These are just two examples, but there is a myriad of opportunities for open data to enhance current non-open data. This calls for a change of mindset in thinking about open data.

[280] Waze (n.d.) available at: https://www.waze.com/en
[281] EENA Help112 Report (2017) available at: https://eena.org/help-112-findings/
[282] EENA Help112 Report (2017) available at: https://eena.org/help-112-findings/

Combining open data with personal, shared, or crowdsourced data is vital for the realisation of further growth of the open data market.

Increasing confidence in data and the re-use possibilities is instrumental to enable value creation. A gap in knowledge and a lack of confidence, prevent re-use. Often, data licensing terms and conditions are not clear or missing, and lead to legal uncertainty. The fear of violating intellectual property rights or the privacy of individuals described in the data is widely present. The recent EU General Data Protection Regulation did not only provide Europe with the legal foundations of modern personal data protection, but also with a solid ground for decision making, i.e. when designing data services that include both personal and open data. With data sharing initiatives, pioneers pave the way for confidence in holistic data re-use initiatives and collaborations. Opening data is vital and a key goal. At the same time, exploring, testing and implementing complementary ways of sharing data is just as instrumental. The mode of sharing is highly dependent on the parties involved and on the purpose. Solutions of value creation with data have to be ethical and sustainable and inspire confidence in their users.

For different challenges we must explore and improve multiple approaches of data re-use that are ethical, sustainable, and fit-for-purpose.

8.3 The 7 key learnings

1. The specification and implementation of high-value datasets as part of the new Open Data Directive is a promising opportunity to address quality & quantity demands of open data.
2. Addressing quality & quantity demands is important, yet not enough to reach the full potential of open data.
3. Open data re-users have to be aware and capable of understanding and leveraging the potential.
4. Open data value creation is part of the wider challenge of skill and process transformation: a lengthy process whose change and impact are not always easy to observe and measure.
5. Sector-specific initiatives and collaboration in and across private and public sector foster value creation.
6. Combining open data with personal, shared, or crowdsourced data is vital for the realisation of further growth of the open data market.
7. For different challenges, we must explore and improve multiple approaches of data re-use that are ethical, sustainable, and fit-for-purpose.

A little longer than four years have passed since the European Data Portal's early economic research "Creating Value through Open Data"[283] was published. The progress that can be observed since, never stops to be a matter of amazement to researchers, practitioners, civil servants and policymakers.

Through the challenges that each participant of this complex and new ecosystem has faced to get to this point, it is relieving to observe the maturity developed, and the awareness of the obstacles still to surmount– as thoroughly described in this report. It is meaningful to be able to put a number to that

[283] "Creating Value through Open Data" (2015) available at: https://www.europeandataportal.eu/sites/default/files/edp_creating_value_through_open_data_0.pdf

progress, measure how long the path we walked together was, and set quantifiable targets that the community can use as benchmarks for the future.

Back in 2015, a quote by EU Vice-President Digital Single Market Andrus Ansip opened the EDP's "Creating Value through Open Data" report: "If I had to express my views about the digital future – that of Europe or indeed, of the whole world - I could do it with one word: data". Today in 2020, that single word still remains the key ingredient to innovation that – in the European Commission President Ursula von der Leyen's words – "can help us to find solutions to societal challenges, from health to farming, from security to manufacturing. In order to release that potential, we have to find our European way, balancing the flow and wide use of data while preserving high privacy, security, safety and ethical standards."[284]

We trust that this report will remain with the readers as a reliable guide as they address those challenges and provide them with methods and tools to measure the impact achieved through the application of open data.

[284] "A Union that strives for more - My agenda for Europe" (2019) available at: https://ec.europa.eu/commission/sites/beta-political/files/political-guidelines-next-commission_en.pdf

APPENDIX: OPEN DATA STORIES FROM EUROPEAN ORGANISATIONS

In this section, we showcase examples of how organisations create economic impact with open data based on their answers in the survey and follow-up interviews.

Actívia Prospect Research & Solutions SL

Company profile

Name: Actíva Prospect Research & Solutions SL
Sector: Professional, scientific & technical activities
Country of submission: Spain
Active markets: Spain
Business relationships: B2G

Description: The organisation offers applied research, consulting, and technical assistance services for the design and implementation of policies and projects within the public and private sector.

URL: https://activaprospect.cat/

Public image of the company

Actíva Prospect tells the public by success stories and gives insights into their daily work. They shared their participation at the @IOpenDataBcn workshop in Spain on LinkedIn (June 2019) - a workshop that focused on the value of open data for public, social, and economic innovation.

In March 2019, employees of Actíva Prospect participated at one of the "Open Data Day 2019" events in Barcelona, which they displayed on Twitter. One of the topics discussed at the event is how open data can be used to analyse public policies. Hence, while the company does not yet showcase open data projects on their own website, they attend workshops and events to engage in the open data community.

The usage of open data

10% of the data used is open data

10% of the open data impacts the revenue

30% of the employee's work (in)directly with open data

Actíva Prospect intends to make **more use** of open data and expects a **little increase** of its impact on their revenue within the next 5 years.

Usage of open data to:

- Enhance services
- Enhance products
- Make informed decisions
- Benchmark and analyse the market

Industries open data is used from are:

Government & public

Further information

- Only 10% of the data you use is open data. If you say that you intend to increase the usage, in what areas do you plan to increase your usage, how do you plan to do this, and what do you see as being the highest potential for your business?
- You say that you only use open data from the government & public. Do you think the overall usage of open data within your company could be increased by considering further industries to extract open data from?

Cerner Corp. Germany

Company profile

Name: Cerner Corp. Germany
Sector: Human health and social work activities
Submission country: Germany
Active markets: 26 countries worldwide
Business relationships: B2B, B2C & B2G
Number of employees: 29,000+

Description: The organisation works to connect people and systems by facilitating solutions that let communities and people engage in their own health.

URL: https://www.cerner.com/

Public image of the company

The organisation is active on different social media channels and releases content connected to open data. For instance, on Twitter the firm wrote a post about open data access in rural areas: Up to 46 million rural Americans have limited access to quality care. With open data access, we can change #RuralHealth for the better (January 2018).

In the blogpost itself, Cerner addresses the concern that open data access is required to further contribute to innovative solutions within health care. They claim that open data access empowers innovators develop tools to assist their stakeholders.

The usage of open data

20%

of the data used is open data

80%

of the employee's work (in) directly with open data

Cerner Corp. intends to **increasingly** make use of open data and expects a **little increase** of its impact on their revenue within the next 5 years.

Usage of open data to:

- Enhance services
- Enhance products
- Make informed decisions
- Increase efficiency

Industries open data is used from are:

- Population & society
- Health
- Government & public

Further information >> next page

Cerner Corp. Germany

Further information

Cerner Corp. Germany uses open data to work with customers to improve the health and care of citizens and patients. The social determinants of health (SDOH) are essential to make healthcare more efficient and better. However, studies have shown that, currently, only 10% of the data points relevant to health management are in a medical context and only 60% of these datasets are related to SDOH.

In the area of SDOH, Cerner Corp. Germany includes data on five areas: economic stability, neighborhood and physical environment, education, social and community context and health and health care.

Economic stability:
- Employment
- Income
- Expenses
- Debt
- Nutrition
- Access to healthy options

Neighbourhood and physical environment:
- Transportation
- Safety
- Parks
- Playgrounds
- Walkability
- Housing

Education:
- Literacy
- Language
- Early childhood education
- Vocational training
- Higher education

Social and community context:
- Social integration
- Community engagement
- Support systems
- Discrimination

Health and health care:
- Health coverage
- Provider availability
- Access to care
- Quality of care

City of Warsaw

Company profile

CITY
OF WARSAW

Name: City of Warsaw
Sector: Public services
Country of submission: Poland
Business relationships: B2B, B2C & B2G

Description: The City of Warsaw is a local self-government unit which executes public tasks of local nature belonging to the competences of gmina and poviat.

URL: https://www.um.warszawa.pl/en

Public image of the company

The City of Warsaw is active on Twitter, Facebook, and Youtube. The municipality does not provide any information in English regarding their view on open data.

The usage of open data

30% of the data used is open data

10% of the revenue is due to open data

10% of the employee's work (in)directly with open data

The City of Warsaw intends to **increasingly** make use of open data and expects a **significant increase** of its usage within the next 5 years.

Usage of open data to:

- Enhance services
- Enhance products
- Make informed decisions
- Benchmark and analyse the market

Industries open data is used from are:

- Agriculture & forestry
- Regions & city
- Education, culture & sport
- Legal & public safety
- Government & public

Further information >> next page

City of Warsaw

Further information

The City of Warsaw has been using open data since 2014 with their Apps4Warsaw project - an initiative that was realised in a partnership with the City Hall, Warsaw University of Technology and ICT commercial company. The core action of this project was to develop applications based on data delivered on the Warsaw Open Data Portal, which was launched as a first step of the project. Since the end of Apps4Warsaw in December 2015, the City of Warsaw has been maintaining and developing their open data portal. Another example of the internal re-use of open data in Warsaw is an international data project called "VaVel", which Warsaw participated in between 2016 and 2018. The goal of this project was to integrate different types of real-time big data sources to develop a prototype of mobile transportation planner.

The City of Warsaw shares more than 200 data archives and collections through the city open data portal. As a result, anyone can easily obtain access to data from official sources on subjects such as transport, education, culture, entertainment, real estate and social projects. Their open data portal serves about 1 million data-base queries per day. The most popular are online datasets (that are continuously being updated), especially those from a public transportation system. Far less exploited - but still popular - are static spatial data that are used as a reference for online datasets, such as bus and tram network or stop locations.

Measuring the social and economic impact of open data is still a big challenge in Warsaw. Nevertheless, the City of Warsaw believes that this impact is positive. Several dozens of applications were built by start-ups, companies and developers based on open data from the city open data portal. Some of those projects achieved commercial success and are still being developed.

The main challenge that the City of Warsaw is facing regarding open data are the silo solutions that make integrating data more complicated. This is more an organisational than a technical issue, where internal data owners (departments and municipal companies) do not always fully understand the necessity of opening data and the value of doing it in a standardised way on city open data portals. Even if this is done, publishing open data is not usually a priority in their agenda. Nevertheless, the City of Warsaw believes that the usage of open data will increase. They are trying to achieve this by:

- Making a relevant redefinition of basics assumptions of city open data portal. Their new data portal (that will replace current website www.api.um.warszawa.pl) will not only be for developers. They are implementing a richer concept of openness to give an open access to city data for those who want to use these datasets. To achieve this, stakeholders were divided into three groups related to digital competences.
    1. Developers are interested in accessing online datasets (M2M formats).
    2. There are people who need data and know how to use them. This group will analyse data if its needed and will visualise, process and share them as well.
    3. There is a major group of people they called "everyman" that do not use much data but will gladly use the tools (based on these data) if it will bring some added values. The city wants to fill their platform with such tools.
- They upload new attractive datasets to the city open data portal. For example, at the end of 2019 they shared online data on current and planned road repairs and investments. They hope that this process will flow faster in the near future.
- They want to give open access to their city data and to datasets provided by external entities (using API standards). Moreover, they want to give information about ICT devices and technologies they use, about conditions of using the city ICT infrastructure and communication standards. They hope that these efforts will bring them closer to the realisation of the local data market. The City of Warsaw believes that creating a local data market will yield wider re-use of open data.
- Working towards increasing an internal re-use of open data. One example is a series of training courses on open data and smart cities that they organised in 2019 as part of the employee self-education programme. These courses will be continued in 2020.

Data Graver

Company profile

Name: Data Graver
Sector: Information Services
Country of submission: the Netherlands
Active markets: the Netherlands
Year of foundation: 2014
Business relationships: B2B

Description: Data Graver is a company that offers its customers IT-related services and possesses expertise in accessing specific data, as well as offering support in the data analysis process.

URL: https://www.datagraver.com/

Public image of the company

The company is active on social media, especially Twitter and Facebook. Some of their posts focus on the implications of open data and on sharing material that helps discussions come into existence.

Furthermore, the organisation shares their personal experience with accessing open data from the Dutch government (2016).

The usage of open data

80% of the data used is open data

50% of the revenue is due to open data

100% of the employee's work (in)directly with open data

Usage of open data to:
- Enhance services
- Forecast scenarios
- Other

Data Graver intends to **increasingly** make use of open data and expects that this trend will **slightly impact** their revenues within the next 5 years.

Industries open data is used from are:

Economy & finance	Education, culture & sport	Energy
Environment	Government & public	Health
Legal & public safety	Population & society	Regions & city
Science & technology	Transport	Other industries

Further information >> next page

Data Graver

Further information

Data Graver is a service company that helps other organisations solve their issues by using their own data or data that is available outside their organisation. 90% of the data from outside their organisation is open data. Examples of open data that Data Graver uses is:

- Population statistics
- Health statistics
- Location information
- General economic indicators (such as PM10, PM2.5, Nitrogen Oxide (NOx))
- Climate data
- Weather data (such as trends and data related to climate change)

In most cases, Data Graver uses open data about singular issues, for example when supplying data on certain health indicators for people living in certain areas. Next to this service, Data Graver provides the media with informative charts and tables on current affairs. Most of these informative charts and tables are also based on open data. In addition, about half of Data Graver's revenue comes directly, or indirectly, from the use of open data.

DesideDatum Data Company SL

Company profile

Name: desideDatum Data Company SL
Sector: Information & communication
Country of submission: Spain
Active markets: Spain
Business relationships: B2B & B2G
Number of employees: 15

Description: desideDatum is a consulting company that provides tools and knowledge for processes optimisation, as well as tools to implement cost and time savings for public administrations and private companies.

URL: https://www.desidedatum.com/en

Public image of the company

desideDatum actively promotes their regular participation at workshops and events that cover the topic of open data on their social media accounts. For example, they posted on Twitter: "@desideDatum will collaborate in a workshop on civic applications and #opendata for public accountability and power monitoring" (February 2019).

Moreover, desideDatum publishes news pieces that cover topics around open data on their website, such as an article about the emergence and progression of open data of the previous 10 years.

The usage of open data

80% of the data used is open data

80% of the revenue is due to open data

100% of the employee's work (in)directly with open data

Usage of open data to:

- Improve services
- Forecast scenarios
- Benchmark and analyse the market
- Increase efficiency

Data Graver intends to **increasingly** make use of open data and expects that this trend will **slightly impact** their revenues within the next 5 years.

Industries open data is used from are:

Agriculture & forestry	Economy & finance	Education, culture & sport
Health	Environment	Government & public
Energy	International issues	Legal & public safety
Population & society	Regions & city	Science & technology
Transport		

Further information >> next page

DesideDatum Data Company SL

Further information

desideDatum is a consulting company that helps public administrations and organisations to overcome the challenges of data society. The company's areas of expertise are open data, transparency, data governance and open government. Regarding open data, desideDatum began using open data since 2011 (formally 2012) and now automates its publication and facilitates its re-use to increase transparency, improve internal efficiency and to create value for the rest of society. In addition, they help organisations improve their internal data governance in order to qualify them for publishing open data. The company uses different types of open data, including market statistics, industrial sector data, geolocation data and more from open data resources.

Currently, open data accounts for approximately 40% of desideDatum's income.

In November 2019, desideDatum published a new report on the re-use of open data in Spain. The report sampled all open data portals in Spain through interviews, by evaluating the different portal's impact and analysing the most popular data topics. Read the report online: https://www.desidedatum.com/informe-opendata/.

DFS Deutsche Flugsicherung GmbH

Company profile

DFS Deutsche Flugsicherung

Name: DFS Deutsche Flugsicherung GmbH
Sector: Transportation & storage
Country of submission: Germany
Active markets: Germany
Business relationships: B2B, B2C & B2G
Number of employees: 5,750 (2018)

Description: DFS monitors air traffic in the German airspace. They have four main control centers and are represented at 16 major international airports in Germany.

URL: https://www.dfs.de/dfs_homepage

Public image of the company

DFS publishes their own magazine: transmission. Topics covered include general news that concern the DFS group, as well as general air traffic industry news.

The usage of open data

DFS intends to **increasingly** make use of open data and expects a **significant increase** of its impact on the revenue within the next 5 years.

Industries open data is used from are:

- Agriculture & forestry
- Population & society
- Environment
- Health

Further information

The DFS is the first major European Air Navigation Service Provider (ANSP) that uses the central database European AIS Database (EAD) for its flight advice. EAD was originally designed to be the database for officially published aeronautical data; however, it is not able to fulfil the intended harmonisation task and is reliant on data published by Member States, which cannot be changed by the EAD.

Concerning open data, DFS is conservative because they are only allowed to use official data for their air traffic management (ATM) systems. Thus, DFS can and will only use open data in the unregulated business of non-aeronautical geodata for the use in Non-Fly Zones (NFZ) for Unmanned Aircraft Systems (UAS). This new field of NFZ services for UAS is the base for DFS's new business Unmanned Traffic Management Solutions (UTM). The service aims to enable a safe and fair integration of UAS in the German airspace. Nevertheless, DFS expects a higher use of open data in the future. For example, they plan to use more open data on the official cadastral information system (ALKIS), especially if it is freely available and if regulated businesses allow it.

Doctrine

Company profile

Doctrine

Name: Doctrine
Sector: Information & communication
Country of submission: France
Active markets: Worldwide
Business relationships: B2B

Description: Doctrine is a set of PHP (general purpose scripting language especially suited for web development) libraries that focus on database storage and object mapping.

URL: www.doctrine.fr

Public image of the company

Doctrine is an open source platform that is freely available under the MIT license. The PHP library was created by a development team that developed it within their spare time. The project is financed donations and partners whose services complement Doctrine Project.

Besides their PHP Open Source library the company offers webinars or offline consulting workshops to help improve the skills and efficiency of external development teams.

The usage of open data

20% of the data used is open data

80% of the revenue is due to open data

100% of the employee's work (in)directly with open data

Doctrine does not intend to make more use of open data and expects a **decrease** of its impact on the company's revenue within the next 5 years.

Usage of open data to:
- Enhance services
- Provide accurate information to legal professionals

Industries open data is used from are:

Economy & finance

Legal & public safety

Further information >> next page

Doctrine

Further information

Doctrine is the first legal information platform in France. Their mission is to open justice by making the process transparent and accessible to everyone. They centralise and hierarchise all legal information (e.g. law, caselaw, doctrine), giving an edge to legal professionals by connecting them with relevant and useful information. More than 2 500 legal professionals already trust Doctrine and every month an estimate of 500 000 people visit Doctrine to find the answers they need to exercise their rights. The platform relies on the availability of legal information, including open data, on French law, French caselaw, EU caselaw, French company register and French lawyer register.

They have been using open data since its creation in 2016 - it is part of their core activity and vision: open justice around the world. Thus, open data is at the core of Doctrine and they aim to provide access to all legal information. They intend to use more open data as it becomes available, provided that it related to the legal field.

Doctrine strongly contributes to the equality of arms principle and the right to a fair trial based on better access to legal information. By reducing asymmetry of information among all the actors of a legal procedure, Doctrine reinforces equality between citizens who do not have the same access to legal information and justice. In addition, Doctrine also participates to significant changes to the legal profession. The complicated access to court decisions goes together with longer hours spent on legal research without even being sure to find the right information for a specific case. In 2019, 88% of lawyers in France failed to access the information they need to defend and counsel their clients. Thanks to Doctrine, legal professionals have an edge on legal information and make better informed decisions when defending their clients, such as www.doctrine.fr/avis. This allows lawyers to reduce costs and reduce litigation costs for French citizens.

Open data increases the revenue of Doctrine. However, as open data in the legal field is processing at a very slow pace, there are some limitations to the re-use of open data.

Eloisa Technologies

Company profile

Eloisa Technologies

Name: Eloisa Technologies
Sector: Administrative & support service activities
Country of submission: Italy
Business relationships: B2B & B2G

Description: Eloisa Technologies provides a virtual assistant offering conversational artificial intelligence to access open data through a natural conversation
.
URL: https://www.eloisatechnologies.com

Public image of the company

The company created a natural language interface for databases and open data from around the world. Eloisa's goal is to create easy accessibility of open data for everyone, including illiterate and disabled people. For example, Eloisa can be asked in a natural conversation for pizza and coffee prices in Milan as long as the corresponding dataset is on file.

The human-like relationships can be established through digital devices, such as computers, tablets, smartphone or smartwatches.

The usage of open data

20% of the data used is open data

10% of the revenue is due to open data

10% of the employee's work (in)directly with open data

Eloisa Technologies is **not sure** whether they intend to make more use of open data, but they expect a **significant increase** of open data impacting their revenue within the next 5 years.

Usage of open data to:
- Enhance services

Industries open data is used from are:

Regions & city

Government & public

Further information

Eloisa Technologies plans to use open data to create value for their idea of Citizen Basic Income. By doing this, they can solve a problem affecting billions of people worldwide - how to have an income to live a decent life in a city of the future that will guarantee housing, food, clothing, medical care and education. Currently, Eloisa Technologies has published all their available solutions on their website. For the aforementioned problem, their proposed solution is artificial intelligence (AI) that will provide mobility services and citizen basic income at no cost for the government.

Eloisa Technologies invites you to read the whole story, to learn more about how they use open data to address the problem of project disclosure and financing, and to contact them: http://www.eloisa.it/RLA/sca/index.htmlU.

ePaństwo Foundation

Company profile

Fundacja ePaństwo

Name: ePaństwo Foundation
Sector: Information & communication
Country of submission: Poland
Active markets: International
Year of foundation: 2010

Description: The organisation's aim is to develop openness and transparency by taking various types of public data and presenting them to society free of charge by using new technologies.

URL: https://epf.org.pl/en/

Public image of the company

As stated on their website, the ePaństwo Foundation provides support to cities, communities, counties and public institutions that want to make their data publicly available for citizens.

The foundation believes that data should help solve problems and that open data is there to facilitate the functioning of cities. However, there is currently no single "best-practice" to open data. Therefore, the foundation helps others with their process to release open data. As of September 2019, they collaborate with a range of municipalities in Poland.

The usage of open data

70% of the data used is open data

10% of the revenue is due to open data

30% of the employee's work (in)directly with open data

The ePaństwo Foundation intends to **increasingly** make use of open data and expects a **significant increase** of its impact on their revenue within the next 5 years.

Usage of open data to:

- Enhance services
- Enhance products
- Make informed decisions
- Benchmark and analyse the market

Industries open data is used from are:

- Economy & finance
- Government & public
- Legal & public safety
- Population & society
- Other industries

Further information

The ePaństwo Foundation moves a huge part of public open data from the "deep web" and makes them visible and accessible in the Internet in new ways by using modern web techniques. The collected public open data is from national registries, such as the National Business Registry, Supreme Audit Office, Public procurements databases, Scientific databases, Patents databases and the National Electoral Commission.

The organisation offers a simple dashboard for their users, allowing them to organise data they care about and to access data easily in the future.

European Environment Agency

Company profile

European Environment Agency

Name: European Environment Agency
Sector: Professional, scientific & technical activities
Country of submission: Denmark
Active markets: European Union (EU), Iceland, Liechtenstein, Norway, Switzerland & Turkey
Year of foundation: 1990
Business relationships: B2C
Number of employees: 200

Description: The European Environment Agency provides independent information on the environment on behalf on the EU.

URL: https://www.eea.europa.eu/da

Public image of the company

The European Environment Agency regularly publishes material that provides independent information on the environment to the public. On YouTube, they present themselves as key provider of environmental knowledge in Europe.

On their website, information can be accessed in form of datasets, interactive maps or graphs. In addition, the European Environment Agency also promotes chosen datasets on their social media pages, such as Facebook, Twitter and Flickr. The open data provided by the European Environment Agency are made accessible on the EU Open Data Portal.

The usage of open data

100% of the data used is open data

100% of the revenue is due to open data

100% of the employee's work (in)directly with open data

Usage of open data to:

- Enhance services
- Make informed decisions
- Make decisions quicker
- Increase efficiency
- Increase the re-use of environmental data
- Actively disseminate environmental data to the public

The European Environment Agency intends to **increasingly** make use of open data and expects a **significant increase** of its usage within the next 5 years.

Industries open data is used from are:

Energy	Environmental	Government & public
Health	Regions & city	Transport

Findtoilet.dk

Company profile

Name: Findtoilet.dk
Sector: Other service activities
Country of submission: Denmark
Active markets: Denmark
Founder: Tine Müller

Description: Findtoilet.dk is a service (online maps and smartphone application) that can be used by municipalities, free of charge, to provide data about public restrooms and make it available to the open community.

URL: http://findtoilet.dk/

Public image of the company

In 2010, findtoilet.dk was among the three winners of the Offentilge Data I Spil (Translation: Open Data in Games) competition in Denmark. Back then, it was only a pilot project. Today, the mobile application service (app) findtoilet.dk has won recognition as an open data app. The app is used as a best practice example for open data in workshops around the world.

The founder, Tine Müller, also presented her project to the European Data Portal where the dataset is available for free. Findtoilet.dk finances itself through competition prize money and donations.

The usage of open data

20%
of the data used is open data

100%
of the employee's work (in) directly with open data

Findtoilet.dk intends to **increasingly** make use of open data and expects a **significant increase** of its usage within the next 5 years.

Usage of open data to:
- Enhance services

Industries open data is used from are:

Government & public

Further information

FindToilet collects open data from local and national governments in Copenhagen and visualises this data on a map in the application to show people where the nearest toilet to their location is and what amenities are available there. It was created to support residents and tourists visiting Copenhagen find the nearest public toilet so that they did not have to visit a cafe or a restaurant, for example.

In November 2019, FindToilet was selected to be in the publication Charter for the Smart City as an example of open data re-use in Denmark.

FindToilet is still expanding and aims to continue until the application includes all 98 municipalities across Denmark. Want to learn more about FindToilet or keen to support the initiative? Visit http://tinemuller.dk/gis-afdelinger/.

Fundación Ignacio Larramendi

Company profile

Name: Fundación Ignacio Larramendi
Sector: Information & communication
Country of submission: Spain
Active markets: Spain & South America
Year of foundation: 1986
Business relationships: B2B & B2C

Description: The organisation's purpose is to contribute to the common interest by offering support for physical or intellectual needs free of charge.

URL: www.larramendi.es/fundacion/home/

Public image of the company

The foundation is active on Twitter and Facebook. On their website they claim to support the development of digital libraries in a globalised context. The organisation's goal is to:

1. Promote charity in social relations, especially for the weak and the disadvantaged.
2. Assess the work of independent institutions as means of optimising resources and invigorating society.
3. Study the historical influence of the Carist movement on Spanish society.
4. Promote non-profit scientific studies and activities. The foundation has set up open libraries that are accessible free of charge to share about 900 polymaths.

The usage of open data

90% of the data used is open data

80% of the revenue is due to open data

90% of the employee's work (in)directly with open data

Fundación Ignacio Larramendi intends to **increasingly** make use of open data and expects a **significant increase** of its impact on their revenue within the next 5 years.

Usage of open data to:

- Enhance services

Industries open data is used from are:

Education, culture & sport

Gemeente Haarlem

Company profile

Gemeente Haarlem

Name: Gemeente Haarlem
Sector: Public services
Country of submission: the Netherlands
Number of employees: 1,000+

Description: Gemeente Haarlem is a municipality located in the North-Holland province in the Netherlands

URL: https://www.haarlem.nl/

Public image of the company

Gemeente Haarlem is active on Facebook, Instagram, Youtube, and Pinterest. The municipality promotes sharing open data on their website in Dutch. Gemeente Haarlem has access to vast amounts of open data, for instance, geographical data, demographical data, and future development plans data. Also, the municipality has a special department that focuses to release open data online and to promote this activity.

The usage of open data

30% of the data used is open data

0% of the revenue is due to open data

10% of the employee's work (in)directly with open data

Usage of open data to:

- Enhance products
- Enhance services
- Make informed decisions
- Make decisions quicker
- Increase efficiency

Gemeente Haarlem intends to **increasingly** make use of open data and expects a **significant increase** of its impact on their revenue within the next 5 years.

Industries open data is used from are:

- Agriculture & forestry
- Government & public
- Population & society
- Regions & cities

Further information

Gemeente Haarlem has been collecting open data since April 2015 to perform its daily tasks. A large part of this data is also interesting for citizens and companies and is available on https://opendata.haarlem.nl. The open data cannot be traced back to people. They represent their open data on the Haarlem website and from there, users can navigate to their open data list where they can find 44 different datasets or see them visualised and styled. It is difficult to objectify the economic impact of open data. Nevertheless, they see a positive economic impact. In addition, since mid-2017, the Operational-Open-Data-Storage (ODDS) has serviced more than one million data requests. Gementee Haarlem uses Geoserve (an open source server for sharing geospatial data) as their OODS. These requests are normally given through the interaction of public servants or other communication channels. Moreover, the majority of requests are generated in applications developed by the organisation and private entrepreneurs.

GetTheData Publishing Limited

Company profile

GetTheData

Name: GetTheData Publishing Limited
Sector: Information & communication
Country of submission: United Kingdom
Active markets: United
Year of foundation: 2017
Business relationships: B2B, B2C & B2G

Description: The company organises open data (only in the UK) into location-based dashboards and point out the sources.

URL: https://www.getthedata.com/

Public image of the company

The datasets created by the company can be accessed via their website or on Twitter, which is the company's main social platform used to engage with the public. In most of their posts they make use of #OpenData.

Besides producing their own open datasets, they take an active part in the topic of open data. On Twitter, they retweet interesting open data sets, for example, from the UK government.

The usage of open data

100% of the data used is open data

100% of the revenue is due to open data

100% of the employee's work (in)directly with open data

Usage of open data to:
- Enhance services
- Follow their mission of exposing and signposting open data

GetTheData intends to **increasingly** make use of open data and expects a **significant increase** of its impact on their revenue within the next 5 years.

Industries open data is used from are:

Education, culture & sport	Energy	Environment
Government & public	Health	Population & society
Regions & city	Transport	

Further information >> next page

GetTheData Publishing Limited

urther information

GetTheData's mission is to expose open data and signpost the source. They re-use open data in ways that help them accomplish this mission. For example, by incorporating the data into dashboards and apps to help discoverability or by combining it with other open data to make it more accessible.

The company mainly use UK government open data and the current focus is on data which is - or can be - geocoded, as they mainly organise data on their website geographically. The company was brought into being with an open data mission, meaning that they have been using open data since the start. The driving force behind the company is to provide an alternative way to discover open data, going beyond the existing data repositories where datasets are marked up with metadata, and exposing the actual data itself in a meaningful context. The limitations around open data that GetTheData have experienced are mainly in the areas of data quality, confusion over embedded rights, and, of course, discoverability.

The company is a commercial venture and revenues are generated by advertising. They are growing steadily and are continually adding more open data to their model as they grow. GetTheData also onward-distribute any data they create or enhance as open data and are seeing that being used by the wider economy.

Examples of GetTheData's open data initiatives are:

- Open Postcode Geo – https://www.getthedata.com/open-postcode-geo
  An initiative that uses British postcodes with additional fields for geo-space applications, including easting, northing, latitude, longitude, postcode area, postcode district, postcode sector, in-code, and ou-tcode.

- Open Postcode Elevation – https://www.getthedata.com/open-postcode-elevation
  An initiative that uses British postcodes with elevation, for example distance above sea level or altitude, in metres.

- Open Flood Risk by Postcode – https://www.getthedata.com/open-flood-risk-by-postcode
  An initiative that uses English postcodes and their risk of flooding.

- Open Units – https://www.getthedata.com/open-units
  An initiative that uses that units of alcohol in branded drinks in a variety of standard servings.

GVA Grimley Limited

Company profile

Name: GVA Grimley Limited
Sector: Public services
Country of submission: Spain
Active markets: 20 countries
Year of foundation: 1974 (group)

Description: GVA Worldwide is a partnership of independently owned professional real estate firms servicing the needs of global occupiers, owners, investors and developers.

URL: https://www.gvaworldwide.com/locations-and-contacts/spain/

Public image of the company

GVA is active on Twitter, but primarily focusses on releasing information about their business activities. GVA Spain does not share their perspective on open data on their website.

The usage of open data

90% of the data used is open data

50% of the revenue is due to open data

0% of the employee's work (in)directly with open data

GVA Spain intends to **increasingly** make use of open data and expects a **significant increase** of its impact on their revenue within the next 5 years.

Usage of open data to:

- Enhance services
- Save costs
- Make informed decisions
- Increase efficiency

Industries open data is used from are:

Government & public

Other industries

Holon Technologies

Company profile

HOLON TECHNOLOGIES

Name: Holon Technologies
Sector: Financial & insurance activities
Country of submission: Estonia
Active markets: Europe
Business relationships: B2B & B2G

Description: Holon Technologies provides telematics technology to integrate different data sources, validate, and analyses data. Their business model focuses on enabling vehicle insurances to improve pricing mechanisms by providing actual driving data.

URL: https://www.holontech.eu/

Public image of the company

Holon Technologies does not actively engage with the public over social media channels. However, on their website they provide the link to the Facebook and LinkedIn accounts of most their employees.

Holon Technologies' business is built upon the usage of data. The company follows the mission to "provide client efficient solutions for data aggregate and analytics".

The usage of open data

20% of the data used is open data

100% of the revenue is due to open data

10% of the employee's work (in)directly with open data

Holon Technologies intends to **increasingly** make use of open data and expects a **significant increase** of its impact on their revenue within the next 5 years.

Usage of open data to:

- Enhance products
- Enhance services
- Save costs
- Make informed decisions
- Make decisions quicker
- Forecast scenarios
- Benchmark and analyse the market
- Analyse competitors
- Increase efficiency

Industries open data is used from are:

Economy & finance	Education, culture & sport	Environment
Health	Legal & public	Population & society
Regions & cities	Transport	

Institute for Market Economics

Company profile

Name: Institute for Market Economics
Sector: Professional, scientific & technical activities
Country of submission: Bulgaria
Active markets: Bulgaria

Description: The Institute for Market Economics is an independent economic policy think tank in Bulgaria. They provide independent assessments and analyses of the government's economic policies.

URL: https://ime.bg/en/

Public image of the company

The results of the analyses and assessments provided by the Institute of Market Economics are available online and freely accessible for the public. For example, they provided a comprehensive report about "German Investments in Bulgaria between 1990 and 2017".

The company does not engage in the public open data discourse nor expresses its view on open data on their website.

The usage of open data

20% of the data used is open data

50% of the revenue is due to open data

50% of the employee's work (in)directly with open data

Usage of open data to:
- Improve services
- Informed decision making
- benchmark and analyse the market
- Increase efficiency

The Institute for Market Economics intends to **increasingly** make use of open data and expects a **significant increase** of its usage within the next 5 years.

Industries open data is used from are:

- Agriculture & forestry
- Economy & finance
- Education, culture & sport
- Government & public
- Legal & public

Lefebvre Sarrut

Company profile

Name: Lefebvre Sarrut
Sector: Professional, scientific & technical activities
Country of submission: France
Business relationships: B2B & B2G
Number of employees: 2,600+

Description: The company aims to provide a service that supports professionals in their effectiveness by developing and extending their knowledge. They concentrate on providing information and training in the legal and tax departments.

URL: https://www.lefebvre-sarrut.eu/

Public image of the company

Lefebvre Sarrut works on several open data issues in collaboration with public bodies to strengthen digital access to public data. It has been announced (April 2019) that Lefebvre Sarrut signed a partnership agreement with the French interministerial directorate in charge of digital affairs and the central government's information and communication systems (DINSIC) to accelerate open public data.

The president of Lefebvre Sarrut, Olivier Campenon said: When it comes to access to legal information, open public data is a citizen's issue. It is a vector of universal rights. It is also one of our Group's missions to facilitate access to the law, in particular through the use of new technologies.

The usage of open data

10% of the data used is open data

10% of the revenue is due to open data

10% of the employee's work (in)directly with open data

Usage of open data to:
- Enhance products
- Enhance services
- Make informed decisions
- Increase efficiency

Lefebvre Sarrut intends to **increasingly** make use of open data and expects a **significant increase** of its usage within the next 5 years.

Industries open data is used from are:

Government & public

Legal & public

Further information >> next page

Lefebvre Sarrut

Further information

Lefebvre Sarrut is a legal publisher that operates in several European countries. Their main product is a legal doctrine where they comment on the law and jurisprudence (primary sources by lawyers and academics). In half of the countries they operate in, the jurisprudence is distributed through open data, including legal data (i.e. law, jurisprudence, some specific administration doctrine and data flow from authorities such as INPI and CNIL). There are also other very specific datasets that they leverage from open data, like collective agreements in labour law (i.e. France) or the tax doctrine from the tax administration (i.e. Belgium). In addition, in some countries they help the administrations to increase the quantity of open data, such as their recent collaboration with the French administration on jurisprudence.

Lefebvre Sarrut is also interested in specific economic datasets that are now available through open data, such as data describing the company activities, both the nature of the activity and financial indicators. These datasets are very valuable when merged with legal data.

In addition, they are also very involved in open source and open data communities. They have been involved in an association called Open Law* le droit ouvert since 2015 and have been taking part in, or running, some open innovation programmes where public administrationd (Supreme Counrt, CNIL, Legifrance, etc.) and private companies discuss and act together to develop open data. In association with Open law, they are discussing how to rationalise their efforts and share work. They hope that these discussions will positively influence politicians to make it easier for stakeholders to act together and share their work. Through open data, stakeholders can limit these operations and focus on added value services for users and clients.

Lefebvre Sarrut has been using open data since it began in France (in 2000). When open data is not available, they need to retrieve the data themselves – sometimes buying it to public administration – and prepare them. All of Lefebvre Sarrut's competitors, and some public authorities, are doing the same. Moreover, they were inspired by some non-related legal open data for new products that they plan to release in the coming months/years. An addition driving force for Lefebvre Sarrut is that without open data, they are likelier to compete with other actors on the quantity of their data (such as the number of legal cases we provide users), which doesn't create added value.

As stated, there are several positive impacts to using open data, including switching from costly licensed subscriptions to free open data flow and additional data to create new services and products. An indirect economic and technological impact is that adding these new data sources and formats can lead to rethinking IT infrastructure, allowing them to leverage more open source-based software, more micro service based, for example. The benefits of these indirect changes are high because these technologies are more modern and better known in the IT world.

Lefebvre Sarrut expects a significant increase along the full implementation of open data in the future. "L'avenir ne sera pas ce qui va arriver mais ce que nous allons faire (the future is not what will happen but what we will do)" Henri Bergson.

Ministry of Sport and Tourism of the Republic of Poland

Company profile

Ministry
of Sport and Tourism
of the Republic of Poland

Name: Ministry of Sport and Tourism of the Republic of Poland
Sector: Public services
Country of submission: Poland
Active markets: Poland
Year of foundation: 2005
Business relationships: B2G

Description: The Ministry of Sport and Tourism of the Republic of Poland deals with matters related to sport and tourism.

URL: https://www.msit.gov.pl/en

Public image of the company

The Ministry of Sport and Tourism of the Republic of Poland is active on Facebook and Twitter. The posts on these channels are framed in Polish and mainly concern Sports and Tourism related matters. For people who do not read Polish it is unclear how the Ministry engages in the area of open data.

The usage of open data

20% of the data used is open data

30% of the revenue is due to open data

10% of the employee's work (in)directly with open data

Usage of open data to:
- Enhance services
- Save costs
- Make informed decisions
- Make quicker decisions
- Benchmark and analyse the market
- Other reasons

Yes, the company intends to **increasingly** make use of open data and expects a **significant increase** of its usage within the next 5 years

Industries open data is used from are:

Communication	Economy & finance	Education, culture & sport
Government & public	Science & technology	Transport

Municipality Nijmegen

Company profile

Name: Municipality Nijmegen
Sector: Public services
Country of origin: the Netherlands
Active markets: the Netherlands
Business relationships: B2B, B2C & B2G
Number of employees: 1.624 (2018)

Description: Municipality Nijmegen is part of the Arnhem-Nijmegen urban region, a metropolitan area with approximately 750,000.

URL: https://www.nijmegen.nl/

Public image of the company

The municipality of Nijmegen promotes open access to its data, as they value easy accessibility and transparency. In addition, they also collect as much relevant information that can benefit other users, such as, journalists, inhabitants, researchers, companies or institutions. Furthermore, the municipality's website allows users to access datasets for different purposes. Lastly, they quote the Dutch government after defining their definition of open data and their vision on the usage of open data.

The usage of open data

Yes, the company intends to **increasingly** make use of open data and expects a **significant increase** of its usage within the next 5 years.

Usage of open data to:
- Enhance products
- Enhance services
- Make informed decision
- Forecast scenarios
- Conduct research

Industries open data is used from are:

Agriculture & forestry	Regions & cities	Science & technology
Transport	Other industries	

Further information >> next page

Municipality Nijmegen

Further information

Municipality Nijmegen started using open data in 2013 via their own data catalogue - opendata.nijmegen.nl. Since then, open data has been an integral part of their organisation and they have been using it in various ways. Their primary use of open data is to improve the services they provide to their citizens. For example, they use their own open datasets to generate maps of Nijmegen, which are publicly available. By using open data, they can clearly display the locations of facilities such as train stations, bus stops, parking garages, etc. Another example that is more seasonally relevant is that they can show which roads are salted in the winter season and at what times the salting takes place. By making this information easily available to the citizens of Nijmegen, their lives can be improved (even if only marginally).

In addition, Municipality Nijmegen also offers several dashboards detailing varying statistics of their city. The primary purpose of these statistical datasets is to improve transparency, to provide insights into what they do, how they do it and how it is measured, and to promote re-usability.

They also use open datasets of other similar Dutch municipalities to gain insights into what they are doing right, or wrong. This information can then be used to improve one or more services that they provide as a governmental organisation. Moreover, they also look at what types of open data are being published and in what way. This is taken into consideration for how and what open datasets Municipality Nijmegen are publishing.

Municipality Nijmegen uses several types of open data. Generally, they use data that focuses on the following:

- Location and other information of facilities within the city of Nijmegen's borders. For example, facilities such as train stations, parking garages, etc. In addition, they use other types of location data, such as the locations of monumental trees in the city area.
- The location of voting booths and data regarding specific voting booths, such as opening times and number of booths per location.
- The usage of electricity, water and gas of the citizens of Nijmegen to make predictions about their usages in the future.
- Data regarding the location, dimension and history of all objects within the borders or Nijmegen.

Municipality Nijmegen is legally obligated to accurately track all real estate within their borders and many of the services they provide are dependent on the accuracy of this data. They also have a legal obligation to use this data for several services municipalities and to provide data on areas such as building permits, the allocation of street names and housing numbers.

Currently, Municipality Nijmegen is seeing a worrying trend where data that used to be publicly available is now being put behind paywalls or is no longer offered altogether. This is a result of more and more organisations realising the value of their data and attaching a price tag on them. For example, government bodies and organisations use data from electric, water and gas companies to make predictions about the expected future use of these resources in the next several years based on population growth and other factors. Now, a part of this data is no longer freely available for others to use. If institutions want to continue to gain insights on the future use of these services, it is likely that they will need to pay money to access the data. As a result, trend legislation is being discussed for specific sectors to publish certain datasets as open data.

Most of the financial gains for Municipality Nijmegen come from statistical analysis performed on their own (open) data. These statistics are used to fine-tune and optimise their internal processes, where possible. The impact this analysis has on their cost and revenues is difficult to quantify in numbers, though it directly affects their operations and operating costs.

National Institute of Public Health and Hygiene

Company profile

Name: National Institute of Public Health and Hygiene
Sector: Professional, scientific & technical activities
Country of origin: Poland
Year of foundation: 1918
Business relationships: B2B, B2C & B2G

Description: National Institute of Public Health and Hygiene is a research institute focusing on public health.

URL: https://www.pzh.gov.pl/

Public image of the company

The Institute is active on Twitter, Facebook, and Youtube. The Institute releases all its information in Polish. For English speakers, therefore, it is unclear how the Institute engages in open data activities.

The usage of open data

The National Institute of Public Health - National Institute of Hygiene intends to **increasingly** make use of open data and expects a **significant increase** of its usage within the next 5 years.

Usage of open data to:

- Enhance services
- Make informed decisions
- Benchmark and analyse the market

Industries open data is used from are:

Environment	Health	Population & society

Open State Foundation

Company profile

Name: Open State Foundation
Sector: Information Services
Country of submission: the Netherlands
Active markets: the Netherlands
Year of foundation: 2006-2008 (merger of two companies)
Business relationships: B2G

Description: The organisation works on digital transparency by opening public information as open data and making it accessible for re-users.

URL: https://openstate.eu/en/

Public image of the company

The foundation actively blogs about open data on their website. For example, the foundation addressed the matter that Dutch public transportation data (OV Data) is still not made available to businesses. They underline that creating open access to this data will not only promote further optimisation of related applications, but also allow researchers and journalists to conduct new studies, for instance.

The foundation shares information on, for example, the concept of digital transparency on Twitter. Furthermore, the foundation organises different meet-ups for the open data community, for instance one where they discuss open algorithms. In addition, the foundation also shares different points of views on open data related practices.

The usage of open data

90% of the data used is open data

100% of the revenue is due to open data

100% of the employee's work (in)directly with open data

Usage of open data to:
- Enhance services
- Enhance Products
- Other

Open State Foundation intends to **increasingly** make use of open data and expects that this trend will **slightly impact** their revenues within the next 5 years.

Industries open data is used from are:

Economy & finance	Environment	Health
Government & public	Legal & public	Population & society
Regions & cities		

Further information >> next page

Open State Foundation

Further information

Open State Foundation (OSF) is a non-governmental organisation (NGO) that promotes digital transparency of the public sector. They do this by advocating that governments release more open data and promote its re-use amongst sympathisers, such as journalists, companies, governments and other NGOs. OSF primarily uses open data on elections, decision-making, public spending and government performance. They also promote the release of data on public transport use (Mobility as a Service (MaaS) and the flow of travellers), company register and UBO-register (which are not freely available yet in the Netherlands). OSF has been using open data since 2007.

Moreover, OSF also uses open data to build platforms, such as openspending.nl or waarismijnstemlokaal.nl or zoek.openraadsinformatie.nl. In addition, they re-use open data to give more insights in data regarding elections, decision-making, public spending and government performance. They advocate that governments should release this information, advise them on how to release the information or improve the data quality, build platforms with the data and promote the re-use of this data at events, such as hackathons and app-challenges. One example is the AccountabilityHack.nl in the Dutch House of Representatives with data on medicine prices and spending on educational money.

OpenCalabria

Company profile

Name: OpenCalabria
Sector: Professional scientific and technical activities
Country of submission: Italy
Active markets: Italy
Year of foundation: 2015

Description: OpenCalabria is an organisation that dedicates space and resources to the preparation of regional economic studies and the development of local development proposals.

URL: http://www.opencalabria.com/

Public image of the company

The organisation is active on both Twitter and Facebook. All information published on their website is in Italian.

The usage of open data

90% of the data used is open data

0% of the revenue is due to open data

80% of the employee's work (in)directly with open data

Usage of open data to:
- Conduct research

Open Calabria intends to **increasingly** make use of open data but believes that it is **doubtful** that this trend will affect revenues within the next 5 years.

Industries open data is used from are:

- Economy & finance
- Government & public
- International issues
- Population & society
- Regions & city
- Science & technology
- Other industries

Further information

OpenCalabria uses open data to carry out scientific reports and analysis based on government data. They re-use open data on demography, macroeconomics variables, public accounts and financial performance of public administration, for example, for scientific purposes and for consultancy. OpenCalabria takes open data from government bodies and public providers of statistical data. By using open data, they increase efficiency, increase productivity and reduce costs.

OpenLitterMap

Company profile

#OpenLitterMap

Name: OpenLitterMap
Sector: Information & communication
Country of submission: Ireland
Active markets: Worldwide
Founder: Seán Lynch

Description: OpenLitterMap provides a web-app that empowers individuals to share open data on plastic pollution anywhere around the world. Therewith, they aim to reduce plastic pollution.

URL: https://openlittermap.com/en

Public image of the company

The company is works with hashtags on social media platforms, such as Facebook or Twitter. Here, they actively inform their followers about open data and its usage. The company writes, for example: Who can download #OpenData? Anyone. What can #OpenData be used for? Anything. How much does it cost to access #OpenData? Nothing (February 2019).

In addition, the founder of OpenLitterMap gave a TEDx presentation about how to challenge plastic pollution with open data which is available on YouTube.

The usage of open data

90% of the data used is open data

10% of the revenue is due to open data

100% of the employee's work (in)directly with open data

Usage of open data to:

- Enhance services
- Enhance products
- Safe costs
- Forecast scenarios
- Benchmark an analyse the market
- Increase efficiency
- Environmental emergency

OpenLitterMap intends to **increasingly** make use of open data and expects a **significant increase** of its usage within the next 5 years.

Industries open data is used from are:

- Environment
- Open Street Map
- Regions & cities
- Science & technology

Samizdat

Company profile

Name: Samizdat
Sector: Information services
Country of submission: Czech Republic
Active markets: Czech Republic
Business relationships: B2B & B2C
Number of employees: 4

Description: An organisation that works with journalistic data, as well as the acquisition, purification, analysis, visualisation and the interpretation of data.

URL: https://samizdat.cz/

Public image of the company

The company is active on social media particularly on Twitter and Facebook. Generally, they deal with processing public data such as election results or historical records of causes of death. They look for hidden connections in, for example, timetables or in the database of towed cars as concisely mentioned on their website.

Furthermore, the organisation is involved in open data activities, as they gave a lecture at the "International Conference of Investigative Journalism" about journalism in relation with data gathering and website data scraping according to their website. The website is in Czech only.

The usage of open data

20% of the data used is open data

40% of the revenue is due to open data

100% of the employee's work (in)directly with open data

Usage of open data to:
- Enhance services
- Save costs
- Make informed decisions
- Forecast scenarios

Samizdat intends to **increasingly** make use of open data and expects a **significant increase** of its impact on their revenue within the next 5 years.

Industries open data is used from are:

Government & public | International issues | Legal & public safety

Further information

Samizdat has been using open data as input for their work as data journalists since 2013. They primarily use public sector data, such as demographic data, transportation infrastructure, environmental data and crime and justice statistics. Samizdat plans to continue making open data more available because data is essential for analyses and projects. In addition, by having readily available open data, companies, researcher and institutions, for example, can mitigate on the cost of searching, requesting, licensing and paying for open data.

The organisation sells content and analytics services to the media so open data can speed up their work and mitigate costs associated to obtaining the open data. Moreover, open data can be used for public relations, advocacy and business analytics. Samizdat hopes that open data will be a standard in the future for publishing government collected data.

SLR Consulting

Company profile

Name: SLR Consulting
Sector: Mining & quarrying
Country of submission: Ireland
Active markets: Europe, Asia-Pacific, North America & Africa
Business relationships: B2B, B2C & B2G
Number of employees: 1,184 (2017)
Revenue: £129,6 million

Description: SLR consulting is an international consultancy firm that focuses on environmental and advisory solutions for private and governmental clients.

URL: https://slrconsulting.com/

Public image of the company

Based on their online presence and own research it is not visible whether SLR Consulting engages in the community of open data. They do not refer to open data on their website and no comments about the participation in open data related workshops or events have been found.

The usage of open data

50% of the data used is open data

40% of the revenue is due to open data

90% of the employee's work (in)directly with open data

Usage of open data to:
- Enhance services
- Make informed decisions
- Forecast scenarios

SLR Consulting intends to **increasingly** make use of open data and expects a **little increase** of its impact on their revenue within the next 5 years.

Industries open data is used from are:

Agriculture & forestry	Energy	Environment
Government & public	Population & society	Transport

Further information >> next page

SLR Consulting

Further information

SLR Consulting uses as much open data that is available, free of cost and relevant to their work as possible. They expect an increase in the availability of open data and a corresponding rise in their (re-)use. This is because open data has a direct economic impact by providing higher quality information to different institutions - i.e. businesses and governments - and reduces the amount of time (and therefore cost) spent trying to replicate the data. In and of itself, data is used for many things, such as Environmental Impact Assessments (EIA).

SLR Consulting produces EIA reports for their clients that include geology, groundwater, soil, hydrology, infrastructure and climate data, for example. Open data makes the process behind these results more efficient and cheaper because the data is already available, meaning that they can prioritise on the quality of the report rather than spend time collecting data.

National Statistical Institute of Bulgaria

Company profile

Name: National Statistical Institute of Bulgaria
Sector: Public administration
Country of origin: Bulgaria
Active markets: Bulgaria
Business relationships: B2B, B2C & B2G

Description: The organisation provides statistical products and services that are produced based on European methodologies and standards.

URL: http://www.nsi.bg/en

Public image of the company

National Statistical Institute of Bulgaria is active on both Facebook and Youtube. The Institute releases most information on their social media accounts in Bulgarian.

The usage of open data

The National Statistical Institute of Bulgaria intends to **increasingly** make use of open data and expects a **significant increase** of its impact on their revenue within the next 5 years .

Usage of open data to:

- Enhance products
- Enhance services
- Save costs
- Make informed decisions
- Make quicker decisions
- Increase efficiency

Industries open data is used from are:

Agriculture & forestry	Education, culture & sport	Energy
Environment	Government & public	Health
Legal & public safety	Population & society	Regions & city
Science & technology	Transport	

UAB CIVITTA

Company profile

Name: UAB CIVITTA
Sector: Other service activities
Country of submission: Romania
Active markets: Estonia, Latvia, Lithuania, Finland, Poland, Ukraine, Romania, Moldova, Russia, Belarus, Serbia, Slovakia, Bulgaria & North Macedonia
Business relationships: B2B & B2G
Number of employees: 303 (2018)

Description: CIVITTA is a management consulting company supporting their clients to become more innovative.

URL: https://civitta.com/

Public image of the company

CIVITTA publishes articles and case studies on their website, Facebook, LinkedIn and Twitter that provide insights into their projects and services.

One of the services they offer, according to their website, concentrates on the help they offer to clients. Namely, to capitalise on the advantages of big data as well as digitising the client's core processes and products. Beyond that, no specific involvement in relation to open data could be identified in their online presence.

The usage of open data

80% of the data used is open data

70% of the revenue is due to open data

70% of the employee's work (in)directly with open data

CIVITTA intends to **increasingly** make use of open data and expects a **little increase** of its usage within the next 5 years.

Usage of open data to:

- Enhance services
- Save costs
- Make informed decisions
- Make quicker decisions
- Forecast scenarios
- Benchmark and analyse the market
- Increase efficiency
- Perform public policy evaluations

Industries open data is used from are:

Economy & finance	Education, culture & sport	Environment
Government & public	Health	Legal & public safety
Population & society	Regions & city	Transport

Further information

Civitta is a big company and each country has their own project. Civitta Romania is using open data primarily in their Public Stream, which consists of approximately 100 employees. The Public Stream provides consultancy services for national, European and international institutions, such as public policy design and evaluation, strategies for local administration and studies on specific topics.

Publications Office
of the European Union

EUROPEAN
DATA PORTAL

Capgemini invent

INTRASOFT INTERNATIONAL

Fraunhofer FOKUS

con terra

sogeti Part of Capgemini

University of Southampton

TIMELEX

52north

the Lisbon council think tank for the 21st century