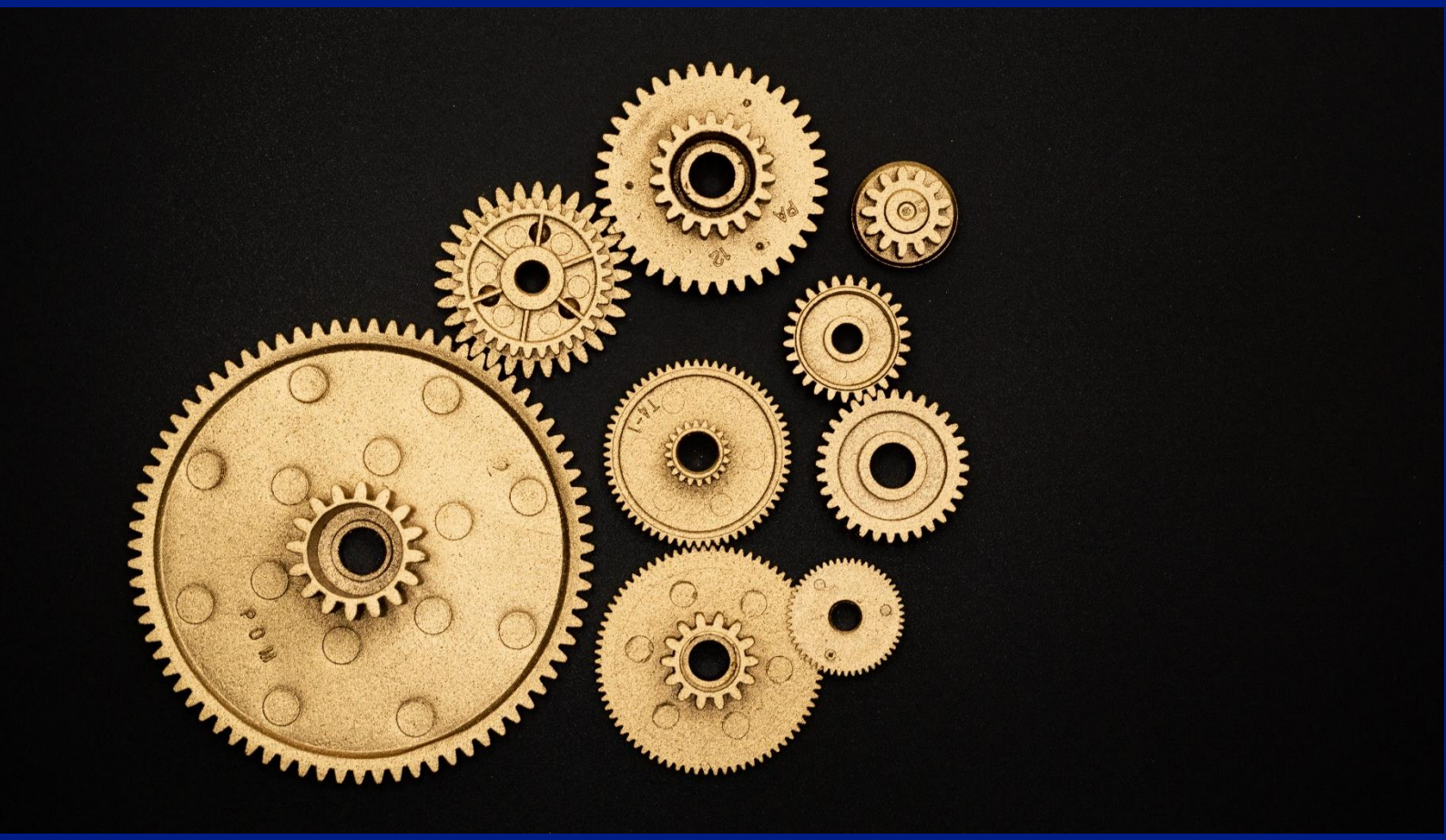
The Use Case Observatory

A 3-year monitoring of 30 reuse cases to understand the economic, governmental, social and environmental impact of open data

Volume I



data.europa.eu

The official portal for European data

Publications Office of the European Union

The Use Case Observatory

A 3-year monitoring of 30 reuse cases to understand the economic, governmental, social and environmental impact of open data

Volume I

Luxembourg, 2022

This study has been prepared as part of data.europa.eu. Data.europa.eu is an initiative of the European Commission. The Publications Office of the European Union is responsible for contract management of data.europa.eu.

For more information about this paper, please contact the following.

European Commission
Directorate-General for Communications Networks, Content and Technology
Unit G.1 Data Policy and Innovation
Email: CNECT-G1@ec.europa.eu

data.europa.eu
Email: info@data.europa.eu

Authors:
Giulia Carsaniga
Jochem Dogger
David Regeczi

Last update: 17 November 2022
https://data.europa.eu/

ISBN 978-92-78-43364-2 doi:10.2830/57907 OA-05-22-401-EN-N

Contents

Executive summary

The use case observatory is a research project that follows 30 reuse cases over the course of 3 years – from 2022 to 2025 – to assess how impact is created with open data, to share challenges and achievements of open data reuse cases and to add to the debate regarding open data impact assessment methodology. This report is the first of three volumes. The second and the third report will be published in 2024 and 2025.

For each of the reuse cases in this report, interviews were held with their developers and clustered into four main impact dimensions: economic, governmental, social and environmental. The interviews on reuse cases with an economic impact show how open data is used to help companies identify promising public procurement tenders and apply for jobs and further professional opportunities. The governmental impact reuse cases prove that open data can increase the transparency of political processes and strengthen democracy. As part of the social impact domain, reuse cases in this report use open data to improve public health and incentivise a more inclusive society. Finally, several reuse cases create an environmental impact by monitoring air quality and supporting the preservation of forests.

Estimating the impact of reuse cases is a complex challenge. The measurement of precise gains for the economy, the government, society or the environment requires a substantial amount of research. Most reuse cases measure their impact with web statistics. Yet, the number of unique visitors or the number of views of a web page only serve as a rough estimate of the impact that is created through open data. Additionally, a challenge in estimating the impact of small and medium-sized enterprise open data reuse cases may arise from the inability to share precise information about their revenue or number of clients, as competitors could use that information to their advantage.

Besides estimating their current impact, the developers of the reuse cases that were interviewed were also asked to forecast the impact they envisage them to have in the next 3 years. The ideas in this respect differ, from only small changes to impressive ambitions. However, a common challenge with regard to the realisation of both small and big plans seems to be receiving appropriate funding to scale up the services offered.

Overall, this first report highlights the power of open data reuse, but also shows the need to further unlock the potential of open data, allowing its reuse to have a clear impact on our economy, government, society and environment. This requires supporting the community of reusers in identifying financial opportunities, but also helping them gain a better understanding of how open data impact is created and measured. While this first report introduces the 30 reuse cases monitored, the second and the third report will focus more on evaluating the progress from one year to another. More emphasis will be placed on ascertaining achievements and challenges in the space of 3 years and extrapolating concrete insights for improving methodologies of open data impact assessments.

1. Introduction

The use case observatory is a research project launched by data.europa.eu, the official portal for European open data, managed by the Publications Office of the European Union. The use case observatory – hereafter referred to as 'the observatory' – contributes to the portal's broader goal of measuring the impact of open data across Europe.

By monitoring 30 reuse cases of open data over 3 years, from 2022 to 2025, the relevance of the observatory is fourfold. Firstly, by following each of the 30 reuse cases, the observatory shows how open data creates impact from an economic, governmental, societal and environmental perspective. Secondly, by analysing the achievements and challenges of reuse cases over a longer period, the observatory allows other open data enthusiasts to learn and increase the overall impact of open data. Thirdly, the long-term monitoring of the reuse cases outlines the entire impact trajectory: from the collection and process of public data to the delivery of the reuse case to the benefit of our economy, government, society and environment. Finally, since measuring the impact of open data remains a challenge, both in Europe and globally, the observatory provides useful insights into the debate on the potential benefits and the several existing approaches of open data impact assessments.

Through the publications of three analytical reports – one in 2022, one in 2024 and the last one in 2025 – the observatory tries to answer the following questions.

- What is the economic, governmental, social and environmental impact of open data for the 30 specific reuse cases collected?
- How important is it to keep track of such reuse cases to understand and foster value creation through open data in Europe?
- What else can be learned from the analysed reuse cases to improve open data measurement and implementation across Europe?

As the first analytical report of the observatory series, this volume sets the stage for the analyses that will be published in 2024 and 2025. Its aim is to introduce the 30 reuse cases in terms of the services they offer, the (open) data they used and the impact they had at the time the report was being written (namely from July to September 2022). The second and third report will expand on the findings reported in this first volume and highlight changes in the development and impact of the same 30 reuse cases reported here. Therefore, the driving questions mentioned above will not find a full answer in this report. Full answers will only be available at the end of the research project and in the publication of the third and final volume in 2025.

The rest of this report is structured as follows. Section 2 focuses on the methodological approach taken for selecting and analysing the 30 reuse cases. Section 3 presents the reuse cases along with the four impact dimensions considered: economic, governmental, social and environmental. Finally, Section 4 sums up the general findings and lessons learned from this first part of the research project and provides an overview of the next steps for the observatory.

2. Methodology

The 30 reuse cases participating in the observatory were selected among a data.europa.eu inventory based on three main sources:

- the open data maturity (ODM) assessment, a landscaping exercise performed every year by data.europa.eu that provides numerous examples of impact creation from EU Member States, the European Free Trade Association (EFTA) and further countries in Europe;
- the EU Datathon, a yearly data-driven competition organised by the Publications Office that offers excellent examples of open data reuse all over Europe;
- the reuse cases available on the data.europa.eu use case repository, which were partially collected by the data.europa.eu team and partially provided via the website's general contact form by data providers in Europe.

Drawing from this inventory, more than 600 reuse cases were identified. To reach the 30 final reuse cases needed, several criteria were used.

Firstly, only reuse cases of applications, platforms or websites developed in Europe were considered. The intention was to keep a fair balance of reuse cases from Member States, possibly including examples from the United Kingdom, EFTA countries and neighbouring countries. In a second step, the focus was placed exclusively on reuse cases belonging to specific sectors in the inventory, with the aim of retaining a good mix of reuse cases having an economic, governmental, social and environmental impact – the four impact dimensions on which the ODM assessment is also based.

- The **economic impact dimension** was defined as including reuse cases in the areas of business creation and/or entrepreneurship, and the (re)skilling of workers.
- The **governmental impact dimension** referred to reuse cases in the areas of e-government support, government transparency and accountability.
- The **social impact dimension** was agreed to encompass reuse cases in the areas of healthcare and well-being, and the integration and fight against inequality in society.
- The **environmental impact dimension** was understood as referring to reuse cases in the areas of environmentally friendly services and energy reduction.

During the selection of the final 30 reuse cases, the following factors were taken into account: whether they had received an award in the last 3 years (e.g. an EU Datathon prize); whether their beneficiaries are women, people from ethnic minorities, people not in education, employment or training, or from any other disadvantaged group in society (e.g. refugees); and whether the reuse cases respond to a crucial challenge of our time, in line with the European Commission's priorities for 2019 to 2024.

This methodological approach allowed 150 reuse cases to be shortlisted, the developers of which were contacted either via email or through social media (e.g. Twitter and LinkedIn) to confirm their availability and interest in taking part in the research. By expressing an interest in more reuse cases than the target number, it was possible to eventually find 30 suitable reuse cases with developers willing to participate.

The developers of the 30 reuse cases were interviewed ([1]) in a semi-structured way via videoconference means, such as Teams or Webex. The interview was conducted in the form of an open

([1]) A complete list of interviewees and respective contacts can be found in Annex I to this report.

discussion based on indicative questions ([2]), shared in advance with the interviewee(s). The questions addressed the purpose of the reuse case, the open data used and the impact the reuse case had at the time and should have in the 3 coming years. Due to time constraints, some interviewees did not participate in the 30-minute interview but provided their answers in written form instead.

Table 1 shows the 30 reuse cases participating in the observatory research project, their country of origin, the impact dimension they were classified in and the source of the use case. Overall, all methodological steps were respected, with only few deviations and adaptations because of a lower response rate than expected due to the summer period.

The final mix of participants represent the EU's northern, southern, eastern and western corners. Yet, it comprises only one reuse case from an EFTA country (Norway), two from the United Kingdom and one from a neighbouring country (Georgia). In terms of impact dimensions, the reuse cases are quite diverse, with 5 economic, 7 governmental, 10 social and 8 environmental examples. While the reuse cases correspond broadly to the definitions of the impact dimensions, it was not always possible to have a perfect match between the areas of reuse cases mentioned above and the reuse cases identified. For example, several reuse cases in the area of healthcare and well-being did participate, but no reuse case in the field of energy reduction was included. Similarly, not all 30 reuse cases have recently won an award, benefit women, people from ethnic minorities or disadvantaged groups in society, or correspond to the Commission's policy priorities. However, at least 25 reuse cases do relate to one or more of the latter criteria ([3]).

([2]) A draft of the questions can be found in Annex II to this report.
([3]) More details on this relation can be found in the respective reuse cases' analyses (Section 3).

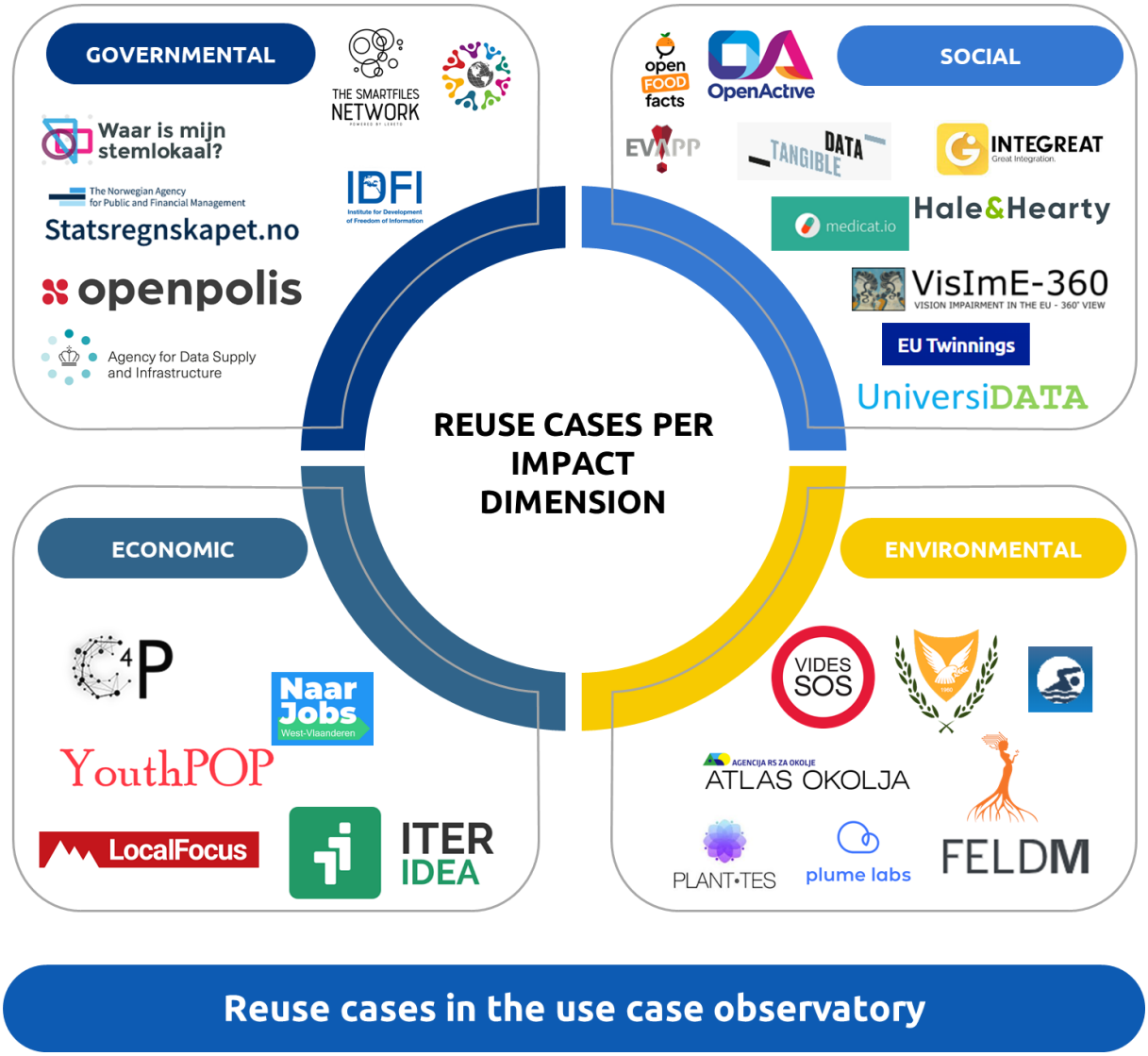
Table 1. Overview of participating reuse cases along with the selection criteria

No	Reuse case name	Country of origin	Impact dimension	Use case source	Further criteria
1	C4P	Belgium	Economic	EU Datathon	An economy that works for people
2	Wonder Wanderlust Women by ITER IDEA	Italy	Economic	EU Datathon	Focuses on women
3	Youth Public Open Procurement (YouthPOP)	Greece	Economic	EU Datathon	Focuses on people not in education, employment or training
4	LocalFocus	The Netherlands	Economic	Use case repository	n/a
5	Naar Jobs in West-Vlaanderen	Belgium	Economic	ODM	An economy that works for people
6	The Smartfiles Network	Austria	Governmental	EU Datathon	A Europe fit for the digital age
7	3D City Model	Denmark	Governmental	Use case repository	n/a
8	Waar is mijn stemlokaal?	The Netherlands	Governmental	Use case repository	A new push for European democracy
9	Openpolis	Italy	Governmental	Use case repository	A new push for European democracy
10	Next Generation Democracy	Denmark	Governmental	EU Datathon	A new push for European democracy
11	Statsregnskapet	Norway	Governmental	Use case repository	A new push for European democracy
12	The Institute for Development of Freedom of Information	Georgia	Governmental	Use case repository	A new push for European democracy
13	Medicatio	France	Social	EU Datathon	n/a
14	UniversiDATALab	Spain	Social	Use case repository	A Europe fit for the digital age
15	ViSimE-360	Italy	Social	EU Datathon	n/a
16	OpenActive	United Kingdom	Social	Use case repository	n/a
17	Tangible Data	Italy	Social	Datathon	A Europe fit for the digital age
18	Hale & Hearty	Ireland	Social	Use case repository	n/a
19	EU Twinnings	United Kingdom	Social	EU Datathon	n/a
20	Open Food Facts	France	Social	ODM	n/a
21	Integreat	Germany	Social	Use case repository	Focuses on migrants and refugees
22	EVapp	Belgium	Social	ODM	n/a
23	Digital Forest Dryads	Romania	Environmental	EU Datathon	A European Green Deal
24	Air Quality Cyprus	Cyprus	Environmental	Use case repository	A Europe fit for the digital age
25	Vides SOS	Latvia	Environmental	ODM	A European Green Deal
26	Planttes	Spain	Environmental	Use case repository	A European Green Deal
27	Atlas Okolja	Slovenia	Environmental	Use case repository	A European Green Deal
28	Plume Labs	France	Environmental	Use case repository	A European Green Deal
29	Baltazar	Croatia	Environmental	ODM	n/a
30	Environ-Mate	Germany	Environmental	EU Datathon	A European Green Deal

3. Reuse cases analyses

This section presents the 30 reuse cases split into clusters according to the four impact dimensions. Figure 1 provides an overview of this clustering into economic, governmental, social and environmental impact. This section also presents a short summary for each of the reuse cases, along with a screenshot of the articles presented later in the report.

Figure 1. Clustering of open data reuse cases per impact dimension



Reuse cases in the use case observatory

Economic impact (five use cases)

1. **C4P (Belgium)** uses a machine-learning algorithm to provide insight into public procurement opportunities from EU institutions and helps clients to find the right tenders and possible consortium partners.
2. **Wonder Wanderlust Women (WWW) (Italy) by ITER IDEA** offers a portal for women between 20 and 35 years old who would like to discover new mobility paths and job opportunities across Europe.
3. **The Youth Public Open Procurement (YouthPOP) web application (Greece)** combines historic Tenders Electronic Daily (TED) data with machine-learning technology to develop an e-tool that empowers young job seekers and entrepreneurs to engage in public procurement processes.
4. **LocalFocus (the Netherlands)** is a platform that provides journalists with a selection of interesting open datasets, easy data visualisation tools and analyses that can be used in articles.
5. **Naar Jobs in West-Vlaanderen – To Jobs in West Flanders – (Belgium)** helps people find jobs close to them and lets users select whether they will travel by bike, car or train. The application also provides information on the transport options provided by employers to new employees.

Governmental impact (seven use cases)

1. **The Smartfiles Network (Austria)** extracts semantics from a Portable Document Format (PDF) text of case-law and relies on network analysis methods to visualise the evolution and interconnectedness of the document with other decisions.
2. **The object-based city model (Denmark)** is a 3D-presentation of Aarhus that can be used for climate adaptation, green conversion, urban planning, land management and much more.
3. **The Waar is mijn stemlokaal? – Where is my polling station? – (the Netherlands)** platform helps users find a suitable polling station close to them. Citizens can also find information about opening times and whether the polling stations are usable for people with certain disabilities.
4. **The non-governmental organisation (NGO) Openpolis (Italy)** gathers, analyses and uses open data for various projects that explain socioeconomic and political dynamics in Italy.
5. **Next Generation Democracy (Denmark)** developed a data-driven solution that enables citizens to follow the work of Members of the European Parliament (MEPs) and engage directly with them.
6. **Statsregnskapet.no (Norway)** is a website that visualises government spending and budgets. Their goal is to facilitate financial transparency and enable the public to easily find information about the spending of resources by government administrations.
7. **The Institute for Development of Freedom of Information (IDFI) (Georgia)** makes governmental data openly available for users on its open data portal and evaluates the transparency of Georgia's government institutions.

Social impact (10 use cases)

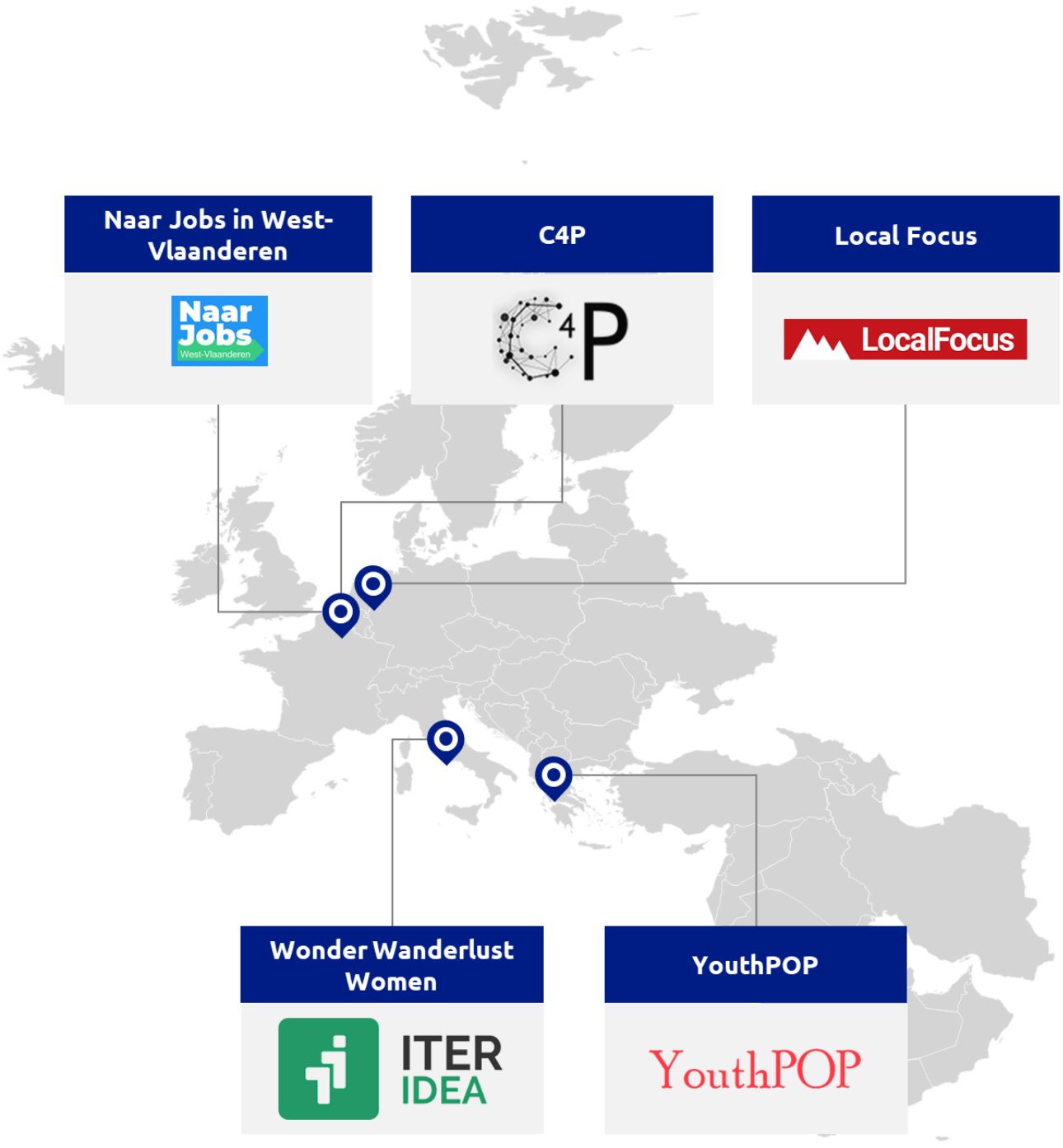
1. **Medicatio (France)** is a platform that publishes data on all drugs available in France to facilitate citizens' access to medical information and allow health professionals to keep track of legal and commercial changes.

2. **UniversiDATALab (Spain)** is a repository of the analytical applications based on the open data published by the six Spanish universities that are part of the UniversiDATA portal. Its aim is to transform the static analyses of a portal's section into dynamic results.
3. **VisImE-360 (Italy)** explains Eurostat's data on visual impairment in a single information web space, helping to allocate resources for medical aid.
4. **OpenActive (United Kingdom)** is a project launched by the Open Data Institute and ukactive to make open data on physical activities in the United Kingdom available and easily bookable.
5. **Tangible data (Italy)** transforms data from its digital context to a physical context by creating data sculptures in the public space, which help people who lack certain digital skills to experience the data.
6. **Hale & Hearty (Ireland)** is a knowledge-based platform and web application created by the Irish government to make information on health and well-being more accessible and to incentivise citizens to live a healthier lifestyle.
7. **EU Twinnings (United Kingdom)** is a website that uses open data from Eurostat to make statistics accessible to a wider audience and show similarities across EU regions.
8. **Open Food Facts (France)** is a large database of food products creating easy-to-understand information about the nutritional value and the environmental impact of food.
9. **Integreat (Germany)** is a digital platform that provides all relevant information in several languages at the municipal level to newly arrived migrants and refugees.
10. **Emergency Volunteer Application (EVapp) (Belgium)** is an application that sends volunteers with first aid diplomas as quickly as possible to someone having a heart attack.

Environmental impact (eight uses cases)

1. **Digital Forest Dryads' application (Romania)** aims to protect forests from illegal deforestation in Europe by combining aerial and multi-spectral satellite imagery.
2. **Air Quality Cyprus (Cyprus)** provides citizens with real-time information about several forms of air pollution. Users can find the data online or choose to be proactively informed about certain substances via the application on their smartphone.
3. **Vides SOS (Latvia)** is an application designed to alert the Latvian State Environmental Service about environmental hazards such as pollution, waste and litter in nature.
4. **Planttes (Spain)** is a citizen science application that informs users about which plants are in bloom and whether they might affect anyone with pollen allergies.
5. **Atlas Okolja – Environmental Atlas – (Slovenia)** presents a map of Slovenia combining a range of different sources, such as noise pollution, air pollution, earthquakes and Natura 2000 areas.
6. **Plume Labs (France)** is a start-up recently acquired by Accuweather that uses open data to forecast air quality globally.
7. **Baltazar (Croatia)** measures water quality on beaches in Croatia. The data is further enriched with information about air temperature, wind speed and beach facilities.
8. **Environ-Mate (Germany)** is an interactive platform to empower children with knowledge about climate change based on scientific data.

3.1. Economic impact dimension



C4P: using open data and artificial intelligence to help organisations find the right public procurement opportunities

C4P in a nutshell.

- **Service:** C4P develops machine-learning algorithms to provide insights into public procurement opportunities from EU institutions and helps clients find the right tenders and consortium partners.
- **Sector:** public procurement.
- **Country of origin:** Belgium.
- **Data sources:** open data from TED, the Commission's financial transparency system, national procurement portals.
- **Number of employees:** 5–10.
- **Website:** c4p.io

C4P was founded in 2017 and won the EU Datathon that same year, it has been growing steadily ever since. The organisation **uses artificial intelligence (AI) to provide insight in public procurement opportunities** from EU institutions and national public sector. These opportunities are published and updated daily on **TED**, the EU's dedicated portal to this end. Although all tenders are publicly available, it can be a challenge to find the opportunity that best suits the competencies of your organisation. To help organisations solve this issue, C4P has created a machine-learning algorithm to support clients in finding the right opportunities.

What services does C4P offer?

C4P makes its own **classification of tenders**, making it easier for users to find the tenders they are looking for. Its machine-learning algorithm uses the title of the tender (around ten words) and the description (around four sentences) to come up with more relevant classifications than are currently offered by TED. To give an example, an EU institution launches a new opportunity about software development services. The title of the procurement is something along the lines of 'Software development service for the European Commission' and the description reads 'The European Commission is looking for a provider of software development in Java'. The functionary who encoded this opportunity might have used the most obvious tags, such as 'IT services', but might have also missed other relevant tags, such as 'software' and 'Java'. C4P's algorithm provides a more accurate classification of tenders.

Additionally to the classification algorithm, C4P also tracks who won previous tenders in a certain field. In that way, C4P's clients can **identify competitors and possible collaborators**. Based on the data of past winners, C4P makes predictions on which organisations are most likely to win a tender that has just been put out on the market. It provides a list to its clients with the top five candidates, meaning the client can subsequently decide to reach out to the candidates to form a consortium.

C4P's clients can be found in all segments and sectors, for example, large enterprises use the organisation to help them find the perfect opportunities and partners. For smaller small and medium-sized enterprises (SMEs), and especially new entrants, C4P's work can be extremely relevant. For

example, the barriers to enter the market for an SME somewhere in Europe could be lowered by having clear information on competition and potential collaborators.

What data does C4P use?

The Tenders Electronic Daily (TED) portal allows users to download the raw data in Extensible Markup Language (XML) format. C4P does a daily download of the data to update its algorithms. Recently, **C4P has also started to expand beyond the TED portal** and also looks at tenders from the United Kingdom's Contracts Finder and Belgium's Public Procurement website. Furthermore, C4P uses open data from the Commission's financial transparency system, in which the Commission services publish their annual accounts (i.e. the commitments done on their financial system). TED and the financial transparency system complement each other, providing a full picture of public procurement: TED shows how procurement was awarded and the financial transparency system shows how it was implemented. This further improves the insights generated by C4P.

C4P would not exist without open data. This goes so far that C4P **only uses open data and no other data sources**. However, after downloading open data, a lot of work still needs to be done. The data needs to be cleaned and structured appropriately before the machine-learning algorithm can do its job. Moreover, data needs to be normalised and standardised to be able to draw the right conclusions. To give an example of such standardisation: company names can be written in many different forms, such as Company XYX, XYZ Co. or XYZ, part of consortium A. Cleaning and standardisation is required to make the data accurate and usable.

How does C4P create an impact?

C4P focuses on the entire **European public procurement market**. Since Europe is in the middle of a twin transition towards a more digital and more sustainable continent, the tender market also contains numerous opportunities in those fields. Hence, the European Green Deal and digitalisation – digital health especially – are the topics that C4P's clients, and businesses in general, are interested in and where C4P creates an impact.

Measuring the precise impact of organisations like C4P is a complex challenge. It is impossible to provide client data and assess whether C4P's clients would have eventually found the same tenders on their own. Still, organisations that work with C4P are often long-term clients and happy with its work. Looking back on the past 5 years since its EU Datathon victory, C4P has grown from a company of two people with a functional prototype into an SME that grows steadily, increases its client portfolios and retains clients. The EU Datathon win was one of the first signs that showed C4P that it was on the right track.

Looking ahead, C4P wants to keep **improving its algorithm in the coming years**. Procurement is used only as a first market, which C4P's team knows very well, but the hope is to **extend products to other markets** as well. Yet, for the years to come, C4P will not leave the procurement market until it has gained a complete picture, for instance, of how to allow its software to read tender specifications (often over 100 pages long) completely and successfully. Recently, C4P has been analysing tender specifications directly, with the aim of moving away from tender notices and having them replaced by AI-generated tender specifications summaries.

Wonder Wanderlust Women: the open data-based portal facilitating women's mobility in Europe

WWW in a nutshell.

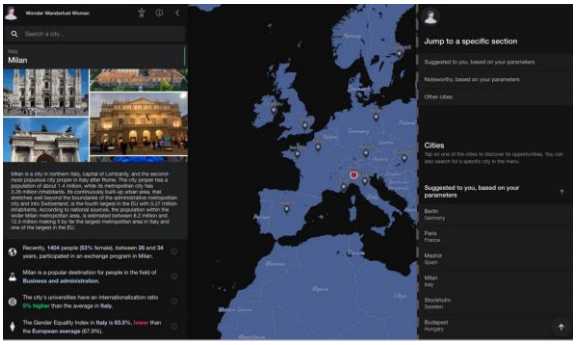
- **Service:** WWW offers a portal for women between 20 and 35 years old who would like to discover new mobility paths and job opportunities across Europe.
- **Sector:** economy, education.
- **Country of origin:** Italy.
- **Data sources:** open data from EU institutions and non-public data from universities and further platforms.
- **Number of employees:** 5.
- **Website:** eu-www.iter-idea.com

WWW is the portal with which the Italian start-up ITER IDEA won the category 'An economy that works for people' during the 2021 EU Datathon. The idea behind the portal is to **encourage the discovery of new mobility paths and opportunities for young people**, **especially women**, across Europe. To do so, WWW uses a large availability of open data to guide women between 20 and 35 years old in the search for the most suitable destination in Europe (where to study, work and live). Since the EU Datathon competition, the WWW team – composed of the three co-founders Sara Baroni, Guido Mazza and Matteo Carbone, and two additional developers – has been expanding its network, collaborating with important entities focused on **gender equality issues**, such as WomenX Impact, and conceiving further data-driven initiatives within the larger context of the ITER IDEA start-up.

What services does WWW offer?

WWW consists of an easy-to-use and engaging web application that visually displays **opportunities and highlights hidden patterns to help young women between 20 and 35 years old** select the most suitable European city and country for their next career and life step. Thanks to the accessibility mode, the user only needs to provide information on their cultural background, country of origin, academic career and spoken languages and the app will suggest all the places that are most likely to match the preferences of the user. Alternatively, WWW also offers the possibility to just wander around the interactive map of Europe and click on specific destinations to find out how many people live there, what is the gender balance in the area and whether jobs or internship opportunities are available.

Figure 1. WWW interactive map and overview of the information provided for the city of Milan



What data does WWW use?

The information provided on WWW comes from different open data sources, including data.europa.eu and related portals. WWW relies, for instance, on datasets about Erasmus mobility, gender equality and WiFi4EU areas. While the role of open data has been crucial in strengthening the experiment phase and creating a concrete base to start developing further, **WWW also uses a number of non-public data** sources retrieved from partner universities, such as IUAV University of Venice, the Numbeo platform (the world's largest crowd-sourced database on cost of living) and different sections of the Erasmus Student Network. The portal – which elaborates over **4 million records to geolocate over 30 000 destinations** – always clearly states the basis of the information shown and provides links to both public and non-public data sources.

Since the data retrieved to build the app was often misspelt, handwritten, translated, incomplete or presented in different layouts, the WWW team had to pinpoint the exact location of heterogeneous datasets in a unique manner. To fulfil this aim, they used **extract, transform and load techniques** on the data and location-technology services powered by **machine-learning capabilities**. Through Amazon Web Services they adopted location, DynamoDB and Amazon Aurora. In addition, the team studied how to improve accuracy by integrating the Google Maps service into their extract, transform and load pipeline. All this was crucial to geo-reference locations, create more value and save over 80 % of input data otherwise lost.

How does WWW create an impact?

WWW targets the lack of support available in Europe for the personal and professional growth of young women. In fact, according to various sources (e.g. the Italian National Institute of Statistics and the Organisation for Economic Co-operation and Development) women between 20 and 35 years old are experiencing the most difficulties in entering the job market. They also represent one of the categories most affected by the economic consequences of the COVID-19 pandemic. Through the WWW platform, ITER IDEA positively contributes to coping with these challenges. In the weeks following the EU Datathon, WWW indeed received **120 requests**, with a subsequent stationary trend of **two to five requests per month**.

While no particular feedback has been collected by the app so far, several collaborations between the WWW co-founders and other institutions have spoken in favour of the growing impact the app could have in the future. For example, the platform has increased an interest in **WomenX Impact**, for which Sara Baroni is now brand ambassador. Together with WomenX Impact, in November 2022, the team will host two panel discussions, one focused on the WWW platform and a second one addressing the topic of 'women in coding'. Moreover, the WWW team is looking into the opportunity of participating in Horizon Europe calls to support a group of organisations in Italy in dealing with violence against women. The idea is to upgrade the WWW app to equip Italian women with new digital support to discover essential support opportunities and define getaway strategies.

Looking ahead, the WWW team will increase its efforts in raising awareness among investors and companies to generate more investment and foster trust in social impact tech solutions such as WWW. In line with the objectives of the European Institute for Gender Equality, and following the publication of a standard for gender equality in organisations in Italy, the team is also undergoing a rebranding phase for WWW to emphasise the powerful meaning of the community they are creating. The establishment of a new and independent enterprise, **UNA Women (United Network Act),** will be an opportunity to actively interact with the project partners, providing for significant additions within the portal.

YouthPOP: the Greek e-tool to empower young job seekers and entrepreneurs in public procurement

YouthPop in a nutshell.

- **Service:** YouthPOP web application combines historic TED data with machine-learning technology to develop an e-tool that empowers young job seekers and entrepreneurs to engage in public procurement processes.
- **Sector:** job market, public procurement.
- **Country of origin:** Greece.
- **Data sources:** open data from TED and Eurostat, and non-public data from European Skills, Competences, Qualifications and Occupations (ESCO).
- **Number of employees:** 3.
- **Website:** youthpop.eu

YouthPOP is an e-tool developed by Michail Maragkakis, Sofia Lousa and Konstantinos Maragkakis, within the context of the 2022 EU Datathon competition. The prototype built by the Greek team aims at **facilitating young job seekers' and entrepreneurs' access to public procurement process** through a user-friendly and interactive web application based on historic open data from TED and Eurostat statistics.

What services does YouthPOP offer?

The YouthPOP project aims to be a hub for young job seekers and business owners – including small and medium-sized entities – where to find different kinds of information on public procurement and related opportunities in Europe. As regards job seekers, a first feature of YouthPOP is to present the **latest trends on open procurement contracts** and **connect these trends with specific types of degrees and skills**, allowing young people to more easily identify their learning needs and must-haves to get their dream jobs. In this regard, the hub also shows young professionals looking to start or change their career the most popular job sectors related to public procurement awarded contracts. Finally, but equally as important, the platform provides its **job-seeking users with detailed information on the quality of life** (i.e. in terms of living conditions and purchasing power) of the Member State they might need to relocate to because of a job or business opportunity.

In reference to its second target group (i.e. young entrepreneurs and business owners) YouthPOP simplifies access to information about:

- the probability of success in given open public procurement contracts;
- potential new clients and partners;
- existing competitors.

The services offered by YouthPOP are presented in a user-friendly and interactive way through colourful bubble graphs, tables and catchy visualisations.

Figure 2. YouthPOP's interface

What data does YouthPOP use?

To develop YouthPOP, the Greek team combined both open and non-open data. More specifically, the **e-tool relies on open datasets from TED and Eurostat**. Moreover, to be able to match public procurement trends and skills, the YouthPOP team also used **free** linked open data, easily accessible from ESCO. ESCO is a Commission project, run by the Directorate-General for Employment, Social Affairs and Inclusion.

After the process of collecting and polishing this data, the YouthPOP team built the application's features via machine-learning techniques, focusing especially on optimising the functionality of the portal, creating inspiring visualisations and enhancing the user experience.

How does YouthPOP create an impact?

At the moment, only a demo version of YouthPOP is available. Therefore, assessing its impact on the target audience is challenging and premature since the team cannot yet monitor the performance of the tool through analytics or users' feedback. However, Michail, Sofia and Konstantinos are planning to integrate such features into the final version of the web application. Moreover, the app's good performance at the 2022 EU Datathon and the positive input received from peers seem to be encouraging signs about the development of the project.

Looking ahead, the YouthPOP team also has a clear set of ambitions to achieve. Firstly, the team members hope to further develop the brand and **increase their solution's visibility** through the EU Datathon competition. Secondly, they would like **to invest in machine and equipment** to better work on the project and be able to work remotely. Thirdly, the team envisages to further develop the application by making it more **accessible for people with disabilities**, in particular individuals with a visual or hearing impairment, by offering the tool in **multiple languages** and by exporting yearly reports and parameters focused on young people. Furthermore, the team would like to explore the possibility of:

- integrating an AI-powered chatbot to guide young people through public procurement;
- aligning the app with existing e-procurement platforms.

Finally, based on the **funding opportunities** explored, including those provided by EU and international institutions, the YouthPOP team might be enlarged to include two more software developers, one chief operations officer and one data analyst.

Ideally, YouthPOP's **ambition is to become an indispensable EU-funded platform**, which is equally helpful for both public and private sectors, and focuses on the needs of young people and possibly beyond.

LocalFocus: enabling journalists to create visualisations with open data sources

LocalFocus in a nutshell.

- **Service:** LocalFocus curates newsworthy data for journalists, provides easy data visualisation tools and performs its own analysis that can be used in journalistic articles.
- **Sector:** media.
- **Country of origin:** the Netherlands.
- **Data sources:** open data from national journalistic sources.
- **Number of employees:** 160 within Algemeen Nederlands Persbureau (ANP), of which 10 work on LocalFocus.
- **Website:** www.localfocus.nl

LocalFocus, which is part of the Dutch press agency ANP, is an **online platform that provides data and data visualisation tools to journalists**. LocalFocus collects and analyses datasets on a wide variety of topics, providing both regional and national newspapers with the ability to transform raw data into visual stories. The platform enables journalists (and non-journalists) to tell data-driven stories, without the need for technical skills.

What services does LocalFocus offer?

LocalFocus is an organisation for data journalism founded in 2012. The company's goal is to help journalists make better use of data, as many journalists still lack the data skills (or simply don't have the time) to find the right datasets, analyse the data and visualise the data for their news pieces. LocalFocus offers three services to make the work of journalists better and easier.

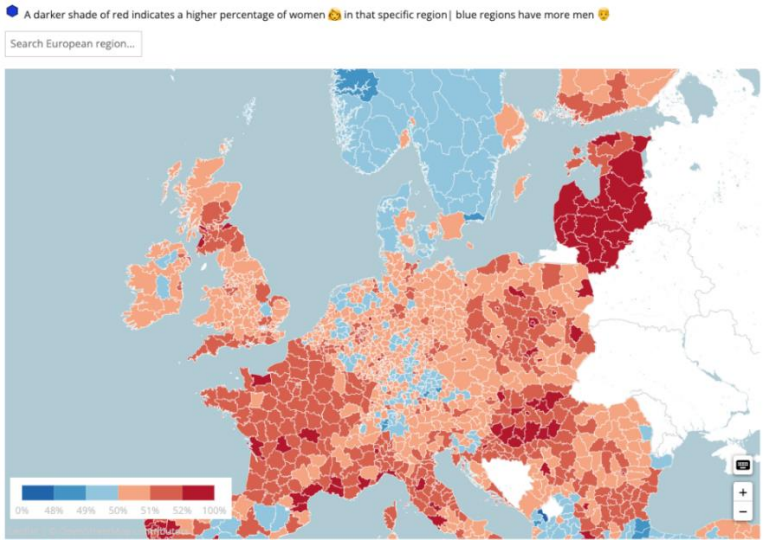
- LocalFocus enriches articles published by the newsroom of ANP with interactive maps and charts based on datasets from different sources. In some cases, the data is also enriched to increase its usability. It chooses the most relevant datasets so that journalists don't have to wade through different websites and platforms.
- LocalFocus has a software-as-a-service platform where data can be uploaded and then visualised. LocalFocus creates a template for its customers, so that journalists and other users can print graphs and maps in the style of their medium. The idea is simple: the user uploads data (e.g. from Excel) and all formatting is done by LocalFocus. This platform is not only used by journalists, but also numerous other users such as analysts in the financial sector who create visualisations for their trend analyses.
- Lastly, **LocalFocus produces between five and eight investigations with visualisations per week**, based on their own analyses. Examples of this are articles on: newly built homes per city; the number of Ukrainian refugees registered per municipality; or simply the percentage of men and women in different regions in Europe (see Figure 3). These analyses are then used by newspapers and other media in their articles.

What data does LocalFocus use?

LocalFocus offers journalists a selection of datasets that they think are relevant. The majority of the data is open data. To name a few of their data sources: the Dutch Central Bureau of Statistics provides data about the number of inhabitants per municipality and the Employees Insurance Agency provides the number of citizens that are currently unemployed. Furthermore, **more niche open datasets are presented by LocalFocus such as datasets on lost pets or the most popular beer brand per municipality**, just as datasets that allow cross-border comparisons are provided by Eurostat.

In some cases, the data is enriched by LocalFocus or a certain clustering is already done (e.g. to group statistics together by region or city). LocalFocus always provides a link to the source data and a step-by-step process description is given whenever data was enriched or analysed by LocalFocus before being published on their platform.

Figure 3. Map of Europe created by LocalFocus showing the percentage of women and men in each region in Europe, made with Eurostat data



How does LocalFocus create an impact?

Thanks to LocalFocus, citizens can be better informed by journalists. Many Dutch media organisations use LocalFocus for their visualisations, and it is also growing in Belgium. The number of times their content is viewed by the media can fluctuate greatly. **At the beginning of the COVID-19 pandemic, their charts and maps were viewed almost 250 million times a month, but nowadays they are usually viewed between 50 and 100 million times a month**. In addition, articles based on research carried out by LocalFocus are copied hundreds of times a month by various national and regional media organisations in the Netherlands.

Feedback is collected through informal means: clients can always call if there are any questions about the platform or the investigations. LocalFocus also helps journalists by actively thinking about what kind of analysis can be done with certain datasets.

At the beginning of 2022, LocalFocus was incorporated into ANP. As part of ANP, LocalFocus wants to look abroad more often and work with international press agencies. Technically, the LocalFocus concept is easily scalable to other countries, but the challenge is to find people with the right expertise who can select and enrich the most relevant datasets available in that country.

Naar Jobs in West-Vlaanderen (To Jobs in West Flanders): finding jobs near you

To Jobs in West Flanders in a nutshell.

- **Service:** To Jobs in West Flanders helps people find jobs near them and lets users select whether they will travel by bike, car or train (or a multi-modal combination). The app also provides information on the transport options provided by employers to new employees.
- **Sector:** economy, job market.
- **Country of origin:** Belgium.
- **Data sources:** open data from national data portals (e.g. vacancy texts, company data).
- **Number of employees:** 5.
- **Website:** naarjobsinwestvlaanderen.be

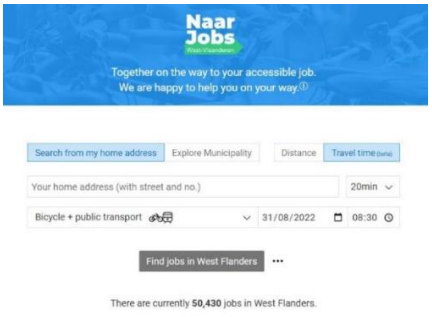
Naar Jobs in West-Vlaanderen (To Jobs in West Flanders) **helps Belgian citizens to find jobs that are suitable to the transport options available to them**. The goal is to help more people find a job and to help companies fill their open spots. Vacancy data is enriched with precise employment sites, enabling social workers to find jobs for unemployed people that suit their transport options.

What services does To Jobs in West Flanders offer?

Nazka Mapps is a Belgian geo-ICT company that created To Jobs in West Flanders to solve traffic poverty together with Mobiel21, who specialise in the social impact transport has on neighbourhoods. Traffic poverty is caused by insufficient travelling options to find a job. Some jobs might be too far away for people or too expensive to get to. Groups that cannot afford a car are especially at risk of traffic poverty. To Jobs in West Flanders is part of a broader programme, which could in theory deliver the 'To Jobs' application to the whole of Flanders (naarjobs.be). The To Jobs in West Flanders website is funded by the Provincial Development Company West Flanders (POMWVL) and the regional department of the Flemish employment service (VDAB) in West Flanders.

To Jobs in West Flanders combines **vacancy data with company data, such as the precise employment sites and transport options (and benefits) they offer to employees**. By combining the address of the job seeker and the precise location and transport options for employees, To Jobs in West Flanders is able to provide an overview of all jobs available within a certain travelling time and with certain modes of transportation (e.g. 60 minutes by public transport).

Figure 4. On To Jobs in West Flanders users can enter their address, transport options and maximum travel time



The app has three main audiences. The first and foremost being social workers that help unemployed people to find a new job. **Social workers actively use the website during one-on-one conversations with people looking for jobs to find a suitable and reachable job together**. The second group are the unemployed people themselves. The reason the main focus lies with social workers and not the people looking for a job directly is that some people might not be as digitally skilled as others or do not speak Dutch, making the website hard to use. The last group are the employers, who can increase the chances of finding competent personnel by sharing accurate data about their location and transport options or benefits they have to offer with To Jobs in West Flanders.

What data does To Jobs in West Flanders use?

The vacancy data that To Jobs in West Flanders uses can be requested via the VDAB open data portal. All vacancies in Flanders are uploaded to this portal. Often, the vacancy provides the address of an employment agency or of one of the main offices of the employer, but not the precise location where the applicant is expected to work. The Belgian Crossroads Bank for Enterprises (VKBO) makes all employment sites per organisation available as open data, and **these employment sites are automatically and algorithmically matched to the vacancies by the 'To Jobs' engine**. When To Jobs in West Flanders is not able to match the vacancy to a precise location in the VKBO data, it creates a new company site and links it to the company data.

To Jobs in West Flanders also uses non-open data. All companies with more than 100 employees need to report to the Belgian government how many parking spots they have for cars, along with the number of places to park bikes and other similar information on transport options. This data is shared with To Jobs in West Flanders to enable it to enrich its data. To Jobs in West Flanders has created an easy online method to add this information for companies with less than 100 employees.

Lastly, open data from Open Street Map and Belgian public transport service providers is used to estimate how much time it takes to reach the employment site. To Jobs has built their own multi-modal routing service combining different transport options.

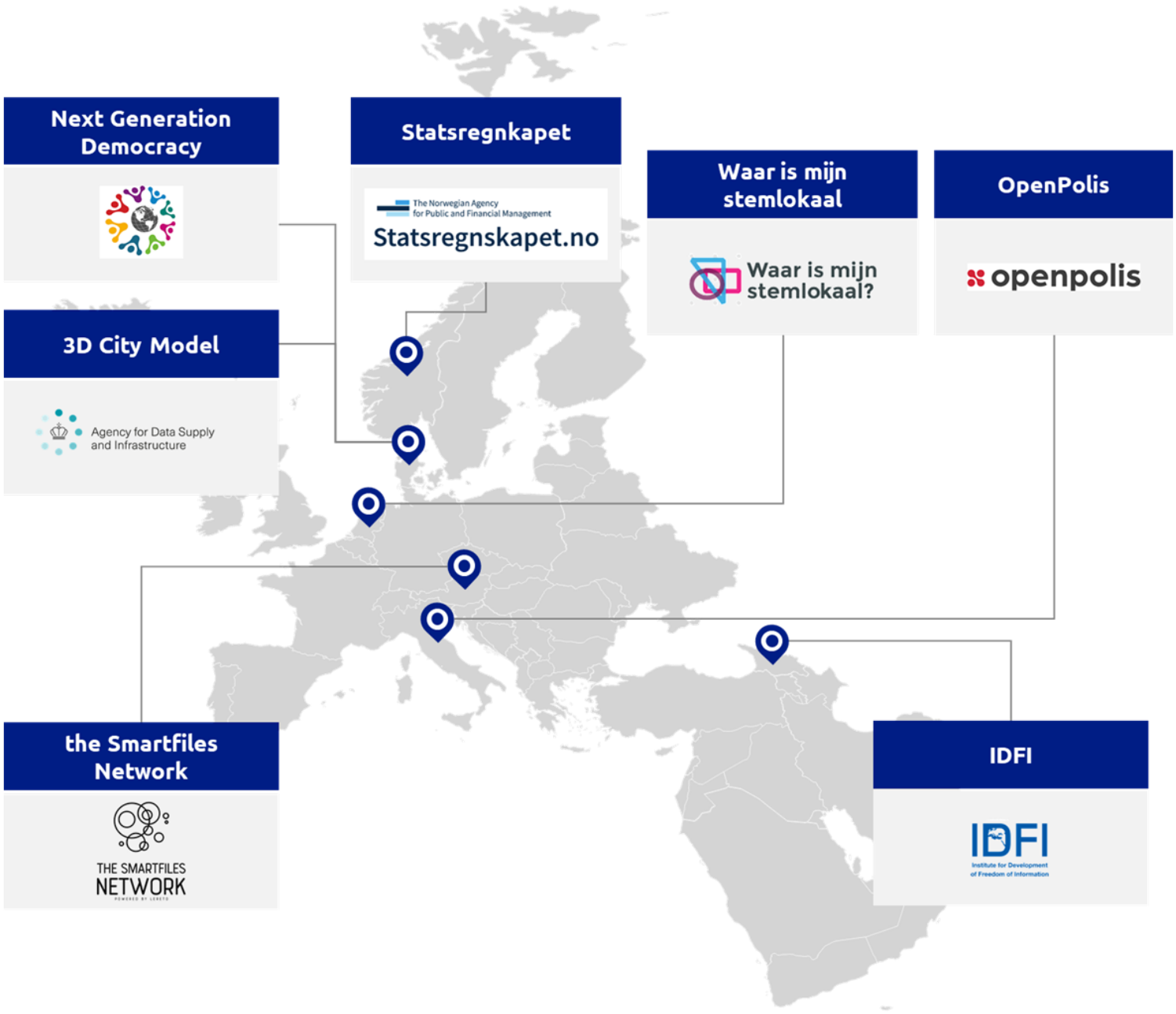
How does To Jobs in West Flanders create an impact?

The website is regularly used by 30 social workers, who are responsible for 10 unemployed people each per day. The website is also used by other non-recurring users. Currently, To Jobs in West Flanders does not know how many people are employed based on the use of its website. **The number of companies registered with its initiative increases by 30 companies per month.**

To Jobs in West Flanders wants to start a campaign to increase awareness of the initiative both with social workers and unemployed people. This campaign was postponed multiple times because of the COVID-19 pandemic. In order to generate more traction, it also wants to expand its reach within Flanders first, then to the rest of Belgium and France and the Netherlands at a later stage.

Lastly, To Jobs in West Flanders is looking to integrate new business models. It has created a data application programming interface (API) where people can download the enriched vacancy data against a compensation. This would enable other organisations to use the enriched data for their own business ideas.

3.2. Governmental impact



The Smartfiles Network: visualising evolution and interconnectedness of EU case-law

The Smartfiles Network in a nutshell.

- **Service:** The Smartfiles Network extracts semantics from a PDF text of case-law and relies on network analysis methods to visualise the evolution and interconnectedness of the document with other decisions.
- **Sector:** EU case-law, government, justice.
- **Country of origin:** Austria.
- **Data sources:** open data from EUR-Lex case-law, EuroVoc and national case-law.
- **Number of employees:** 4.
- **Website:** smartfiles.lereto.at

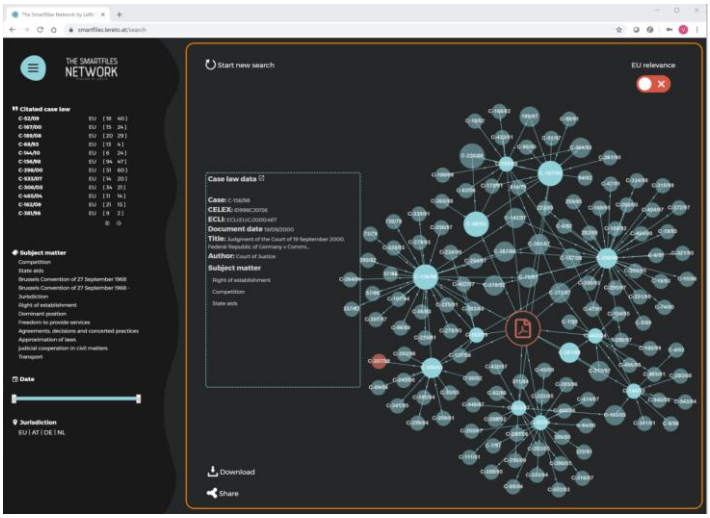
The Smartfiles Network is part of LeReTo, a legal-tech start-up founded in Vienna (Austria) in 2014. Since then, the LeReTo start-up has won several awards, most recently in 2019 at the EU Datathon. On this occasion, Veronika Haberler and Peter Melicharek co-founded the solution 'The Smartfiles Network', with which they triumphed in the category 'Innovative ideas through EU open data'. Aware of the fact that legal professionals deal with thousands of pages of PDF texts every day, Veronika and Peter developed The Smartfiles Network to **simplify legal research and more innovatively visualise case-law.**

What service does The Smartfiles Network offer?

The Smartfiles Network relies on an algorithm that is able to extract the text from a EU case-law PDF and, through recognition logics, find the relevant information about the European case-law identifier (ECLI) that is used to classify court decisions in the EU. This identifier is necessary for the algorithm to calculate the **in- and out-degree and display these in a real-time citations map**. The in-degree is the value representing how many other case-law decisions cite the given document, or in other words how relevant the given document is within the EU network. The out-degree, on the other hand, refers to how well-founded the case-law is (i.e. how many citations are found within a judgement). The real-time citations map based on these values – the first being variable and the second being constant over time – shows the evolution and interconnectedness of the analysed PDF document, which The Smartfiles Network further **enriches with database links** that enable citations to be clicked on. The citations map can be easily downloaded and shared via social media.

In a nutshell, The Smartfiles Network makes PDFs on EU case-law interactive and more easily navigable for **legal professionals, but more generally for anyone working with research on EU case-law** and EU case-law decisions. In fact, in the last 2 years, publishers, content providers and EU institutions such as the European Parliament have shown an increased interest for The Smartfiles Network as a new way of interacting with legal text and legislative procedures. The Smartfiles Network team was ultimately invited to the Court of Justice of the European Union in January 2020 to present the project and the idea of a document-based search.

Figure 5. Citations' map



What data does The Smartfiles Network use?

The Smartfiles Network would not be able to work and **calculate in- and out-degrees without open data**. Moreover, in the legal domain, it is necessary to work with open data across Europe. To build the algorithm, The Smartfiles Network team relied on EU level open data from EUR-Lex case-law and EuroVoc, cite-scraping the PDF files and further metadata (ECLI and case number, uniform resource locator (URL), subject matters, dates). As regards national data, the team also gathered information from the Austrian public legal database, the RIS (Rechtsinformationssystem des Bundes) and conducted a pre-check of data from the Rechtsprechung im Internet in Germany and De Rechtspraak in the Netherlands. The team had to build their own database for the in- and out-degree calculation and clean the data for citations to account for the fact that it is not common everywhere to use the ECLI when citing EU law. For example, in Austria, but also in Germany, it happens more often to use the case number.

How does The Smartfiles Network create an impact?

Since its victory at the 2019 EU Datathon, The Smartfiles Network has not been tracking its number of users, but there are different qualitative proofs of interest demonstrating that the tool has gained popularity among various stakeholders. With interactive visualisations, The Smartfiles Network is not only **revolutionising the world of PDFs but is most importantly making legal texts easier to grasp**, simplifying the work of legal professionals, researchers and policymakers, while potentially also **improving the access of EU citizens to key court decisions**.

While the COVID-19 pandemic represented a slowdown in the development of The Smartfiles Network, the ambition of the team is to foster the impact of this open data reuse case by **integrating the measure of relevance in LeReTo's core technology**. Moreover, the team also has the intention to refine the algorithm by **aligning its definitions of in-degree and out-degree with the research and Marc van Opijnen's recommendations**, legal informatics adviser at UBR|Kennis- en Exploitatiecentrum Officiële Overheidspublicaties (the Publications Office of the Netherlands) and leading mind behind the ECLI. Finally, the team will work on an upgrade of The Smartfiles Network website application.

3D City Model: making future decisions based on a realistic 3D representation of Aarhus

3D City Model in a nutshell.

- **Service:** the object-based city model is a 3D presentation of Aarhus that can be used for climate adaptation, green conversion, urban planning, land management and much more.
- **Sector:** government.
- **Country of origin:** Denmark.
- **Data sources:** open geodata from Geo Denmark.
- **Number of employees:** 3.
- **Website:** dataforsyningen.dk

Lifelike digital 3D models of urban environments have long been part of computer games, but there is much greater potential in the use of 3D city models. The Danish Agency for Data Supply and Infrastructure and the municipality of Aarhus have developed a **prototype of the entire municipality of Aarhus to investigate what is possible with a modern object-based 3D city model** (the visualisation contains one tile of 1 km^2).

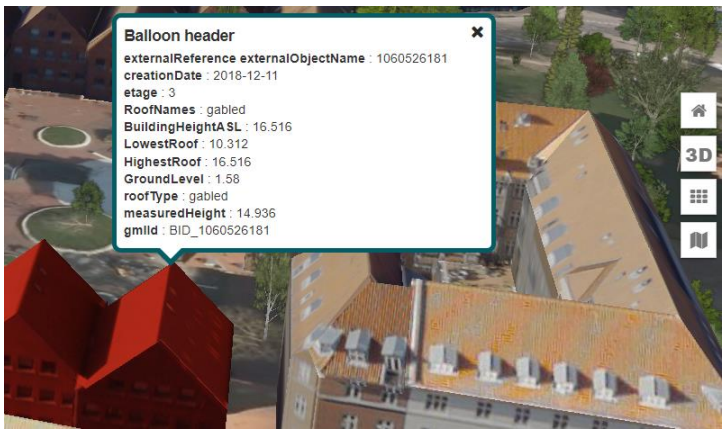
What services does 3D City Model offer?

The use of 3D city models as a communication tool between citizens and public administrations is obvious, for example, in consultation processes on new urban development projects.

The 3D city model can be used as a tool of analysis, visualisation and communication for multiple purposes. For instance, the model can provide **climate adaptation simulations of flooding risk after heavy rainfall** in the city or help find the right location for clean energy sources such as windmills and solar panels. Because of the object-based data, the model shows which roofs obtain the most sunlight and which locations are optimal for windmills.

The Danish Agency for Data Supply and Infrastructure has developed an **object-based version of a 3D city model for Aarhus**. Object-based modelling means that each building constitutes a separate object with a unique identification (ID) associated with it. Thus, when you hover over a building with your mouse in a 3D map you can find specific information on a given building, such as the height, the number of floors and the roof type (e.g. flat, gabled or shed). The figure below shows the information that pops up for each building.

Figure 6. Screenshot taken from 3D City Model with information about the selected building (in red)



3D City Model also provides benefits to the private sector, such as information about facade renovation, window cleaning and many real estate-related matters.

What data does 3D City Model use?

The city model pilot in Aarhus uses geo data from Geo Denmark as its main source and this data is openly available. Data from Denmark Footprints is used for the energy consumption of buildings and data from Pointcloud is used to get the height of a building. These data sources are connected via the object IDs of the buildings on the map.

The Danish Agency for Data Supply and Efficiency (SDFE) has outsourced the visualisation of the 3D map to Burec (Building Reconstruction) and Terrasolid, who use the (mostly open) data to generate the buildings on the map.

A challenge for 3D City Model can be that the **models are based on disparate data of varying ages, data formats and standards**. This can complicate analyses that cross municipal boundaries, for example, if the model is to be used to map the location of a new railway line or where and how much a river overflows its banks during periods of heavy rainfall.

How does 3D City Model create an impact?

The long-term goal for the SDFE is to create a nation-wide 3D map, but as of yet a lot of work still needs to be done. Today, **about 25 % of the country's municipalities use local 3D city models**. Interest in 3D city models among municipalities is high but setting up and maintaining 3D city models is costly. The Aarhus 3D city model is one of the few that is well maintained and open to users.

Feedback is collected via two ways, the first being the feedback option on the website of the prototype where users regular submit their comments. The SDFE also receives feedback from the municipalities they are working for. This enables them to tweak the model for both citizens and local governments.

The current impact of the Aarhus 3D city model is limited, given that the project is currently still a prototype. **Upscaling both in size and in impact largely depends on the funding** the team at SDFE is able to get. The funding can be provided by the government (as part of Denmark's digital strategy), but also options for private-sector funding are kept open.

Waar is mijn stemlokaal? (Where is my polling station?): making voting easy and inclusive for everyone

'Where is my polling station?' in a nutshell.

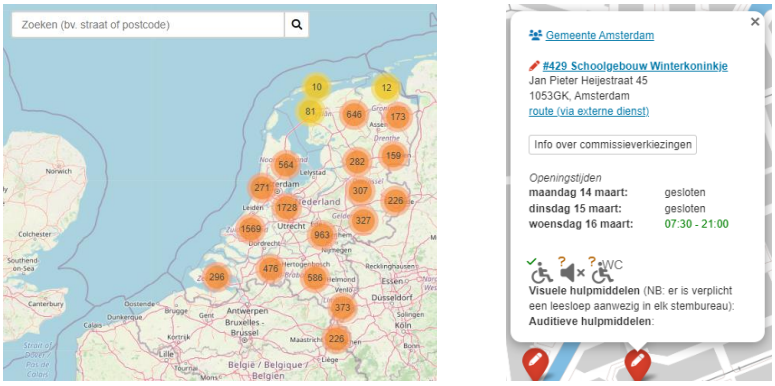
- **Service:** the 'Where is my polling station?' platform helps users find a suitable polling station near them. Citizens can also find information about opening times and whether the polling stations are usable for people with certain disabilities.
- **Sector:** non-profit, government.
- **Country of origin:** the Netherlands.
- **Data sources:** open data from national polling stations.
- **Number of employees:** 5–10.
- **Website:** waarismijnstemlokaal.nl

The Waar is mijn stemlokaal? (Where is my polling station?) platform was created and is maintained by the Open State Foundation, a non-profit organisation whose mission is to open up public sector information as open data and making it accessible for reusers, believing that it strengthens democracy and creates substantial civic and economic value. Other open data initiatives from the Open State Foundation focus on lobbying transparency and public spending. The 'Where is my polling station?' website **lets citizens locate their nearest voting booth** on a map and allows the polling stations to be filtered based on several accessibility criteria.

What services does 'Where is my polling station?' offer?

Over 13 million Dutch citizens were able to vote in the 2022 municipal elections. In 334 municipalities, voters can go to the polls. But where can they vote? The 'Where is my polling station?' application **helps citizens find a polling station that suits their needs**. The website lets users search for a polling station based on distance from their home and opening times. Moreover, the website offers additional information for people with specific needs, such as citizens who use a wheelchair or people with a visual or hearing impairment. The website informs citizens, for instance, about whether the polling station is accessible by wheelchair, whether there are accessible toilets and about the acoustics. Polling stations in the Netherlands are obliged by law to have magnifying glasses available.

Figure 7. On the left – overview of all polling stations in the Netherlands. On the right – information about one specific polling station



The main audience of the 'Where is my polling station?' website are, of course, citizens looking for a place to vote, but this is not the only group of users. The website also serves as the **main platform to find information about the votes per polling station**. In other words: if you want to understand the Dutch elections at a granular level per polling station, then the 'Where is my polling station?' website has the most accurate information. The platform also offers all data in bulk. Journalists therefore also regularly use the bulk data from 'Where is my polling station?' to support their news articles.

What data does 'Where is my polling station?' use?

The Open State Foundation does not use non-open data and publishes all the data they use. The website first started by collecting information about polling stations themselves from the websites of local governments, which suffered from many data inequalities. To solve this issue, they created a harmonised data standard and asked all municipalities to upload their data on the website in the standardised format. Currently, 82 % of municipalities follow the standard created by the Open State Foundation, which means that for 18 % of the municipalities the data still needs to be collected manually.

The 'Where is my polling station?' website uses open cadastral data, geo data from Open Street Map and specific information about buildings with bag ID (base registry for addresses and buildings) to further complement the data from the municipalities.

How does 'Where is my polling station?' create an impact?

During the 2022 elections, the website was visited by **500 000 unique users**, who remained on the website for an average of about 1 minute. For the 2021 parliamentary elections this number was even higher, **reaching 700 000 unique visitors**.

The data is also regularly used by the media. The Dutch Broadcast Foundation, for instance, drew attention to some polling stations that opened earlier, to give people who are at risk from COVID-19 the chance to vote in a safer environment. A large Dutch newspaper also used the data to visualise the vote share of political parties per municipality, meaning that everyone could find out how the votes cast in their polling station were distributed.

The website is constantly **being updated and improved based on user feedback**. Especially during elections, where they receive a lot of feedback from users on their website. They also set up **research groups to better understand the needs of people with disabilities**. For the near future, the Open State Foundation's goal is to secure long-term funding to keep improving the completeness of the information and the user friendliness of the website, so that municipalities can continue to provide data for users to access.

Openpolis: the Italian foundation that narrates political power with open data

Openpolis in a nutshell.

- **Service:** Openpolis gathers, analyses and uses open data to explain political power in Italy.
- **Sector:** non-profit, government.
- **Country of origin:** Italy.
- **Data sources:** open and non-public data from national government entities.
- **Number of employees:** 15.
- **Website:** www.openpolis.it

What started in the 2000s as a simple association, the Italian foundation Openpolis was officially launched in 2018 with the aim of making data available for reuse to the greatest possible number of citizens. Since then, Openpolis has not only focused on the **technical aspect of providing open data, but also on narrating this data**, affirming itself in the area of data journalism as well. Openpolis has accumulated a consistent number of datasets that are regularly updated and used for a series of different projects, all focused on **making socioeconomic and political transformations in Italy more transparent**.

What services does Openpolis offer?

Openpolis offers a variety of different services to its audience, which ranges from students and journalists to institutions and administrators.

- Web applications based on open data, such as **Open Parlamento**, which monitors almost in real-time the activities of the Italian parliament based on data retrieved from the Italian House of Representatives and the Senate of the Republic.
- Online platforms, such as **Centri d'Italia**, developed in collaboration with Action Aid. The platform maps out – through a user-friendly and freely accessible website built on data from the Italian Ministry of the Interior – all reception centres for refugees and asylum seekers in Italy.
- Initiatives like **Open bilanci**, which offers citizens direct access to the balance sheets of Italian municipalities.
- Personalised information, as in the case of **Open PNNR**. Based on the data of the Italian Ministry of the Economy, the platform informs subscribers about the latest developments in the implementation of the Italian recovery and resilience plan through regular email updates and catchy visualisations.
- Further projects, including **Mappe di potere**, which gathers data related to Italian politicians and connects it with data on central administrations (e.g. ministries, constitutional entities) and data on the Italian economy (e.g. public and private companies' data). The aim is to trace properties and shares and understand the multiple roles that politicians and administrators might have.

Each of these initiatives and datasets contribute to Openpolis' content creation published on the Openpolis.it magazine and on the related newsletter. This year, Openpolis also founded its own

membership programme, with the aim of increasing its audience, further engaging the existing community activities in Openpolis and contributing to the **financial and political independence** of the foundation.

Figure 8. Openpolis' range of activities

What data does Openpolis use?

The idea of Openpolis is rooted in the belief that data is key for our society and needs to be used for the greater good (i.e. to increase people's awareness and participation in democratic processes and decision-making). Unsurprisingly, Openpolis has long been **lobbying Italian government entities and administrations to make more of their produced and used data open**. In fact, besides Open Bilanci, which uses data that has been historically open (as required by the Italian law), many other initiatives of the foundation make use of data that was often – but not always – obtained through **Freedom of Information Act procedures**. Hence, a big effort is made not only to clean, refine and analyse data and create content with it, but also to exert pressure on administrators in Italian ministries, governmental agencies and municipalities to make this data openly accessible.

How does Openpolis create an impact?

By contributing to the publication of public institutions' data in Italy, Openpolis has a clear impact on the state of open data in the country. By using this open data, the foundation also seems to be reaching an increasing number of users, in alignment with its goal of making data available and reusable for as many citizens as possible. For example, the Open PNRR project counts **1 600 subscribers**, including institutions, associations, companies, regional and local administrations, and single citizens. The newsletter, on the other hand, has **between 80 000 and 90 000 subscribers**.

Beyond the Italian reality, Openpolis is closely following what is happening at the EU level. To contribute to the EU goal of overcoming data silos and fostering data sharing, **Openpolis is working on gathering in one single point of access the data of public institutions**, along with that of private and public companies in Italy, especially at the local level. As set out in the Data Governance Act, Openpolis would like to pursue the logic of **data altruism**, unlocking the potential of open data not only through a top-down, but also a bottom-up approach.

Next Generation Democracy: the Danish NGO that uses open data to improve democracy

Next Generation Democracy in a nutshell.

- **Service:** Next Generation Democracy offers activities and data-driven solutions to improve democracy and raise awareness on a use of technology that benefits all.
- **Sector:** non-profit, government.
- **Country of origin:** Denmark.
- **Data sources:** open data from EU institutions.
- **Number of employees:** 2–6.
- **Website:** nextgenerationdemocracy.com

Next Generation Democracy is an NGO based in Denmark whose goal is to **raise awareness and drive change by leveraging data and new technologies to achieve a democracy that works for all**. To achieve this aim, the NGO organises workshops and interviews with global thought leaders and identifies like-minded organisations and opportunities to use technology for the greater good. Next Generation Democracy's initiatives includes the **Future Readiness Index**, the solution with which Michael B. Jensen (co-founder of Next Generation Democracy) and Kristian T. Madsen (member of Next Generation Democracy) won the 2020 EU Datathon category 'A new push for European democracy'. The prototype they developed enables citizens to follow the work of MEPs and engage directly with them.

What services does Next Generation Democracy offer?

How can technology help avoid some of the challenges of democratic processes? With this question in mind, Next Generation Democracy performs different activities, among which include keeping a global **democracy-technology database**, with all great initiatives, ideas and organisations within the democracy and technology ecosystem. Moreover, the NGO organises **workshop sessions on topics at the crossroads of technology and politics**, such as AI for democracy or predicting policing, and interviews global thought leaders in the field. Finally, Next Generation Democracy identifies opportunities to **leverage technology for the improvement of society** and tests them through simple experiments.

One of these experiments is **the Future Readiness Index**. Based on data about MEPs' amendments, the web application **categorises politicians along 450 different topics and ranks their work according to the future relevance of the topics treated**. For example, green energy or AI legislation are topics that would make an MEP receive a high score, while coal power plants would receive a low score. The index does not consider the view of the given politician on the specific topic, but rather ranks them based on how much time the politician spent on a certain topic: the larger the amount spent on green energy, the higher the score of that politician ([4]). Conversely, the more time is used for coal power plants, the lower the scoring. When using the Future Readiness Index, users are able to see the ranking by political group, individual MEP or country. For each MEP, information on nationality, political affiliation and their 'fingerprint' is provided. The fingerprint icon on the web application shows a word

([4]) Information on how the index is constructed is open for feedback and can be found in the section ´above` of the web application.

cloud that summarises the main topics of interest for the given MEP. By clicking on the Twitter icon, on the other hand, users are directed to the MEP's social media page and by clicking on the email icon, they can directly write to them. Below this information, the web application offers users the possibility to read through the various activities that the consulted MEP is involved in, with their respective score.

Figure 9. Interface of the Future Readiness Index and an example of an MEP's fingerprint

What data does Next Generation Democracy use?

To build the Future Readiness Index, the Next Generation Democracy team used **open data from the European Parliament and, more precisely, data on amendments in which MEPs** were involved in. Since they needed data that could indicate individual behavioural tendencies, the team also considered using voting data. Yet, given MEPs often vote on blocks (i.e. along political affiliations) the team decided to exclude this kind of data. Another data source that was taken into account is recordings of legislative procedures and transcriptions of discussions that could allow a sentimental analysis to be carried out. This data, however, was too complicated to properly use within the time frame of the EU Datathon competition. To perform a textual analysis and web-scraping of the amendments' data based on the word list of EU topics, Next Generation Democracy firstly considered going directly to the source and using APIs from the Parliament. Yet, because of the inaccessibility and bad quality of some of this data, the team decided to appoint an intermediary.

How does Next Generation Democracy create an impact?

The 2020 EU Datathon was a huge motivator for Next Generation Democracy. Yet, **scaling their Future Readiness Index following the win was not as easy** as expected. Due to the loss of data and related analytics on the side of the intermediary, the team was unable to update the index prototype for more than 6 months. When the data was finally retrieved again, the team was able to reconstruct the demo version of the index (i.e. the one based on data from 2020), but the time wasted and the few financial possibilities did not allow them to advance much further. The inconvenience with the data led to a momentum loss and to a cut in team members. This challenged Next Generation Democracy's ambition to scale up the solution and engage more users.

Despite this, the Danish NGO has continued to have an impact on democracy and society through its other services and keeps on looking positively to the future of the index. In fact, besides the success gathered at the EU Datathon, the ranking solution has been **attracting attention at various other events** – most recently at the Copenhagen Democracy Summit – and among MEPs themselves, who could use the tool to more easily find and collaborate with like-minded colleagues.

Looking ahead, the Next Generation Democracy team would like to **raise further financial resources** – through a partner or a foundation – to build up a team around the Future Readiness Index and scale up the prototype by increasing user engagement. They also envisage to **upgrade the index** to perform more tailored rankings and to extend the tool at local level.

Statsregnskapet.no: helping citizens understand how the government uses taxes and public revenues

Statsregnskapet.no in a nutshell.

- **Service:** statsregnskapet.no is a website that visualises government income and spending to facilitate transparency and enable the public to find information about the central government accounts.
- **Sector:** government.
- **Country of origin:** Norway.
- **Data sources:** open data from government agencies (e.g. income, expenditures/costs, work-year / full-time equivalents).
- **Number of employees:** five people within the Government Agency for Public and Financial Management (DFØ), two full-time equivalent (FTE) work part-time on statsregnskapet.no.
- **Website:** statsregnskapet.no

Transparency is a prerequisite for accountability. Given the size of government income and spending, transparency regarding where the money comes from and where it goes is essential for a well-functioning democracy. Statsregnskapet.no provides an **overview of the Norwegian government's income and spending** in an easy-to-understand dashboard.

What services does statsregnskapet.no offer?

Statsregnskapet.no is run by the DFØ and **presents accounting data, mostly cash accounting data, but also accruals accounting data for some government agencies** (companies), appropriations data and FTE data in several views. The user can view graphic presentations of income and spending at government and ministry (department) level, along with income and expenditures or costs and FTE information for government agencies. The user can also compare the development of appropriations and expenditures data monthly, per item, chapter, programme category and programme area. The 'About state (central government) accounts' page includes stories, definitions and other relevant information to enable users to understand and make good use of information on the website. A comparison page makes it very easy to compare accounting data, key financial indicators and several indicators for different expenditures or costs per FTE at agency level over time, and across agencies.

Figure 10. Screenshot taken from the (translated) statsregnskapet.no website

Statsregnskapet.no has a very broad audience. Simply put, anyone with an interest in Norwegian public administration is the target audience, which includes ordinary citizens, politicians, the media, people in academia, special interest groups, the Norwegian Office of the Auditor General, consultants, students, leaders and employees in government agencies, along with the private sector. In short: **statsregnskapet.no is a unique source of knowledge for anyone interested in Norwegian public administration.**

What data does statsregnskapet.no use?

Most data on statsregnskapet.no is open data. The annual reports from government agencies, including annual accounts and FTE data, are published on the respective websites of the government agencies. Appropriations data is available via the government's own web page (regjeringen.no).

The monthly accounting data presented on statsregnskapet.no is originally not open data. **DFØ receives the accounting data reported by the government agencies on a monthly basis and performs some validation controls before making the data accessible** as open data on statsregnskapet.no and on dfo.no for others to use.

Cleansing and structuring of data is done in procedures available through database / Structured Query Language tools. Statsregnskapet.no does not manipulate or change any data, but some key figures are put together from more than one data source (e.g. key indicators such as salary per FTE).

How does statsregnskapet.no create an impact?

DFØ monitors the use of statsregnskapet.no on a regular basis. However, since the website is an entirely open solution without any form of login required, visitors are not individually identifiable. Using a web analytics program, DFØ measures the number of visits (**15 500 annual average 2018–2021**), the number of new and previous visits, average visit duration, the number of actions per visit and a range of other indicators that provide information about how and how often the website is being used. DFØ uses this as input for marketing initiatives, competence measures and reporting.

DFØ seeks to promote statsregnskapet.no on a regular basis, typically when new annual data is presented on the website, when significant improvements are implemented or new types of data are made available for users. In recent years, to create awareness of and stimulate interest in the solution, DFØ has been more active and experimented with a mix of **blogs, short videos, articles on social media, newsletters to government agencies and conducting webinars**. These activities compliment the courses and workshops available for journalists, government agencies, the Norwegian Office of the Auditor General and others. Course feedback and user surveys indicate that statsregnskapet.no is both useful and user friendly.

In the coming years, DFØ will continue to develop the website and proceed with efforts to **increase the use and utility of statsregnskapet.no.** Moreover, in August 2021, the Norwegian Ministry of Finance decided that accruals accounting principles should become mandatory and implemented for all government agencies within a 5-year period. Therefore, a main goal the coming years will be to **enable reception and presentation of accrual accounting data for all government agencies** in statsregnskapet.no.

The Institute for Development of Freedom of Information: making closed or inaccessible public sector data available in formats that are easy to use

IDFI in a nutshell.

- **Service:** IDFI makes governmental data openly available for users on its open data portal, and evaluates the transparency of Georgia's government institutions.
- **Sector:** government.
- **Country of origin:** Georgia.
- **Data sources:** open and non-public data from government bodies.
- **Number of employees:** 30 (of which seven work on activities related to access to open data and supporting journalists).
- **Website:** idfi.ge

IDFI **obtains large amounts of data from public institutions in PDF format and publishes it in machine-readable formats** on an online data portal. Examples of data published by IDFI are access to legal information, court decisions and open parliamentary data. IDFI also provides yearly rankings of the transparency of government institutions in Georgia.

What services does IDFI offer?

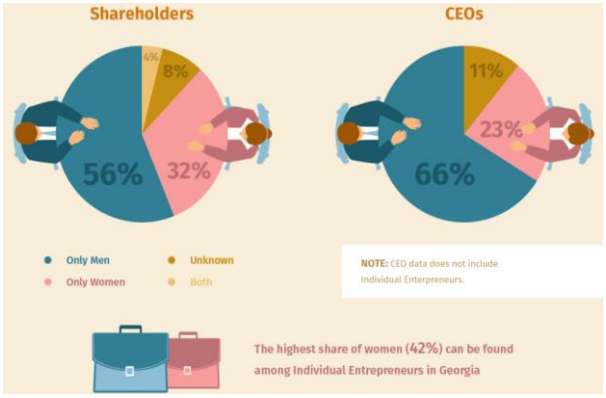
IDFI was established in 2009 and has three main strategic goals:

- to encourage an informed public by expanding the public's access to information;
- to improve the quality of democratic governance by building open, accountable and responsive national and local governments;
- to invest in a more equitable society by promoting policies and practices demonstrated to build political, social and economic equity and improve overall well-being.

Especially for the first two goals, open data and opening up data is a prerequisite for success.

With the help of international donors, such as the Visegrad Fund and the Netherlands, **IDFI created its own open data portal (datalab.ge)**. On this portal users can find around 1 400 governmental datasets, almost all available in Excel and comma-separated values (CSV) files. Moreover, the website offers a toolkit to help users use the data and the data visualisations created by the IDFI team. IDFI also organises annual contests for open data users to raise awareness about the opportunities that open data offers.

Figure 11. Data visualisation created by IDFI based on open data from its portal



IDFI also publishes a yearly report about the transparency of government institutions, in which they check the compliancy of the government with freedom of information (FOI) requests. The data they collect for this report is subsequently made available to the public. **The report provides a transparency ranking of government institutions**, partly to celebrate ministries that are doing well, but mostly to name and shame those that regularly fail to comply with the request. Moreover, IDFI also advocates for better transparency legislation.

Other activities of the IDFI include advocating for open data standards, to make open data easy to use for everyone in Georgia. Furthermore, IDFI helps stakeholders retrieve information from the government that is not yet available, and has a legal aid project for journalists, which helps them to go to court when FOI requests are not granted.

What data does IDFI use?

IDFI uses many open data sources from the government in its open data portal. However, in many instances this data is still in PDF or Word format and **IDFI first needs to transform the data into Excel or CSV files to make reuse easier**. IDFI uses an open-source software for transforming data from Word/PDF to Excel/CSV file when possible, but often the data still needs to be manually added into Excel. Updating the data after the government publishes new data also has to be done manually.

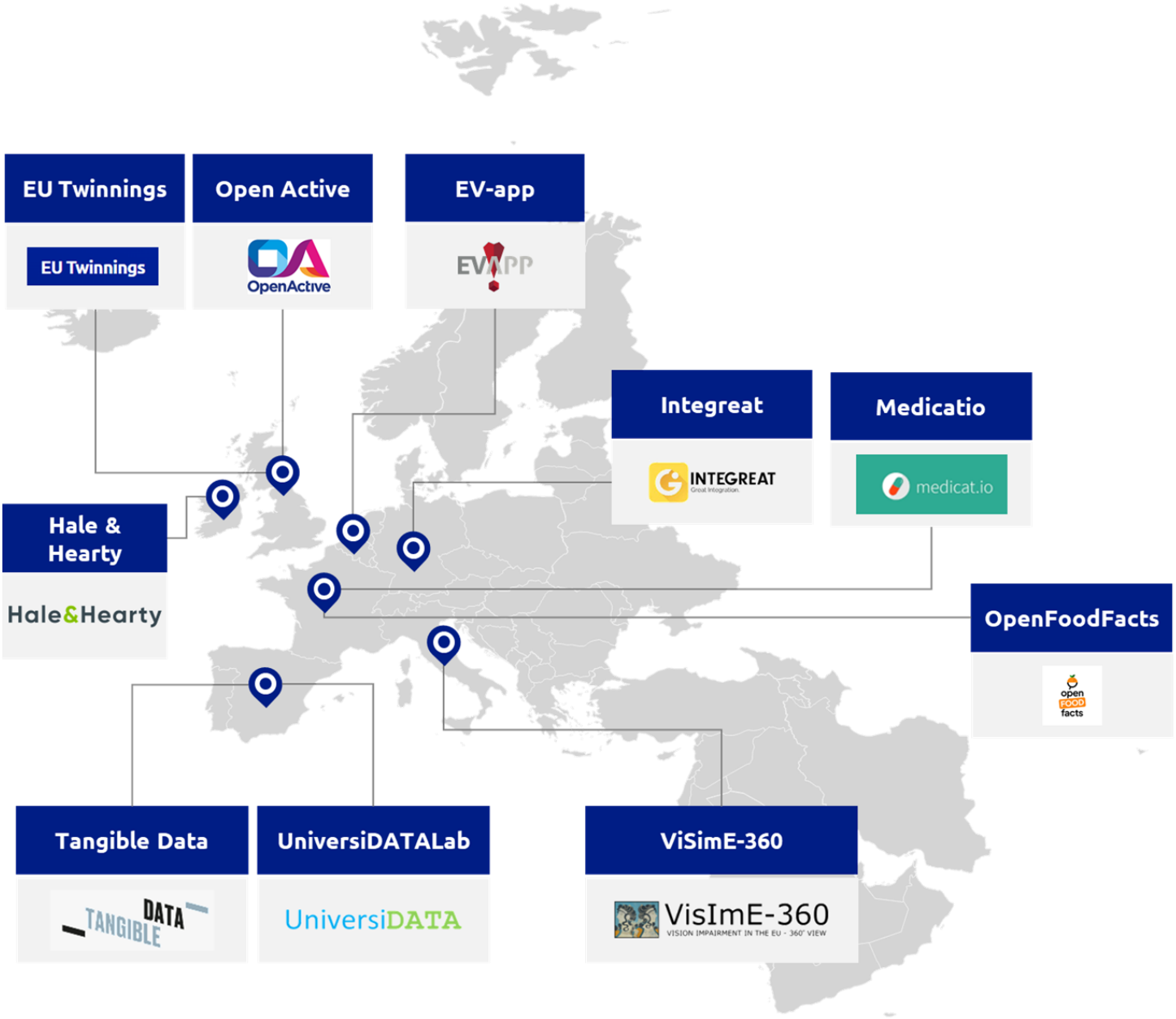
IDFI makes a lot of data available with FOI requests. This data is also often delivered to them in PDF format and then needs to be changed into a machine-readable format.

How does IDFI create an impact?

IDFI's annual report gets lots of media coverage each year and the outcomes are uses by multiple civil society initiatives. Moreover, the top-performing government institutions in the report also write their own articles celebrating their score. **Transparency in the Georgian government improved substantially since IDFI started reporting it** in its annual monitor. Although this improvement cannot be contributed fully to IDFI, it is likely that the reports played a role.

IDFI's open data portal has 50 users on a regular day, but recently had over 1 000 visitors per day during its data for change challenge, of which the goal was to encourage the use of open data to prepare analytical articles. 15 articles were published as a result of the contest. For the coming years, IDFI aims to expand the number of datasets available for reuse on the open data portal and to explore whether it can publish data that is currently held by companies, such as data from hotels to analyse tourism in Georgia.

3.3. Social impact



Medicatio: the open-data-driven app that provides user-friendly information on medicines

Medicatio in a nutshell.

- **Service:** Medicatio publishes data on all available drugs in France to facilitate citizens' access to medical information and allow health professionals to keep track of legal and commercial changes.
- **Sector:** health.
- **Country of origin:** France.
- **Data sources:** open data from the European Medicines Agency and national health agencies.
- **Number of employees:** 1–2.
- **Website:** medicat.io

Medicatio was launched in 2015 from an idea conceived by the engineer Willy Duville, joined later by PhD candidate Miwon Seo, who aspired to **mind the informative gap between citizens and governmental authorities releasing information about drugs**. By providing user-friendly information on medicines, including usage and drug interactions, the Medicatio platform impressed the EU Datathon jury of 2018 and won the challenge 'EU open data – For more innovation in Europe'. Since then, Medicatio has continued relying on open data to keep its users updated on the 15 000+ approved drugs in France and in the EU.

What services does Medicatio offer?

Medicatio applies cutting-edge technology to offer a user-friendly website, where citizens can find free, neutral, up-to-date and easy-to-understand **information on all approved drugs in France and in the EU**. The service provided by Medicatio allows people to overcome the sometimes-biased profit intentions of commercial organisations and the heavy jargon of doctors and pharmacists.

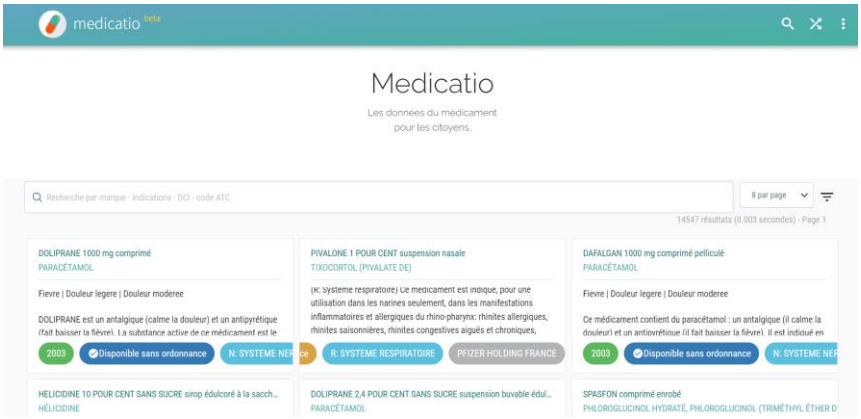
At the same time, with frequent updates of the numerous approved drugs in France and in the EU, Medicatio also represents a reliable source of information for health professionals, who wish to check the latest changes in the medicines prescribed to patients.

While the main beneficiaries of Medicatio is the general public, the team is considering involving more doctors and health professionals in the uploading and updating of data and recommendations on the specific drugs displayed on the platform. In fact, for the moment, Medicatio only uses officially published open data by medical authorities in France and the EU.

What data does Medicatio use?

The main source of information for Medicatio are the **open datasets** of the Agence national de sécurité du médicament et des produits de santé (National Agency for the Safety of Medicines and health Products) and the big database on reimbursement of prescribed drugs 'Assurance maladie' (health insurance) of the Haute Autorité de Santé (French National Authority for Health), which provides evaluations, notes and useful documents on drugs for health professionals. A second important source of information for Medicatio comes from the European Medicines Agency.

Figure 12. Research banner on Medicatio



Currently, the Medicatio team still adds some datasets themselves, but the idea is to make the uploading and updating process as automated as possible. The team has prepared a pipeline of downloaded data from the above sources. However, after downloading this data needs to be cleaned to ensure the highest quality of the given information. This is particularly important for data on medicines given that most information retrieved is still hand typed in a verbose style by health providers and might contain typos – the Medicatio team have already reported a few mistakes to the responsible authorities. Following quality assurance, all data downloaded and cleaned from the Agence National de Sécurité du Médicament, the Haute Autorité de Santé and the European Medicines Agency is combined in one single database through a framework based on GraphQL, an open-source data query and manipulation language for APIs.

How does Medicatio create an impact?

Since winning the EU Datathon in 2018, Medicatio has been growing steadily reaching more and more people, especially in the last months of 2022. In June 2022, the number of active users of the website ranged between **1 000 and 2 000 users per day**. This greater outreach is certainly due to Medicatio's intuitive features and accessibility, which is further amplified by the fact that the team improved the web semantics to make the website easily findable on Google. In other words, when someone searches for the name of a specific drug and some additional key words, they are automatically suggested to visit Medicatio.

The intention to more proactively **involve health professionals in updating the information** on Medicatio will also allow the platform to have a greater impact on the healthcare sector in France and in the EU, for which providing accessible health data is a priority.

In the future, Medicatio would like to **distinguish itself even more from other companies** and platforms providing a similar service and from the sources and software used by them. With this aim, Willy Duville, from the Medicatio team, is **collaborating with the Laboratoire d'informatique médicale et d'Ingénieurie des Connaissances en e-Santé** (French research laboratory focusing on medical informatics and e-health research) to review the existing algorithms and better identify drug-on-drug interactions. The first research paper on this collaboration should be published in 2023, following which the algorithm will be updated on the platform as well.

UniversiDATALab: showcasing the potential of open data for Spanish universities and beyond

UniversiDATALab in a nutshell.

- **Service:** UniversiDATALab is a repository of the analytical applications based on the open data published by the six Spanish universities that are part of the UniversiDATA portal. Its aim is to transform the static analyses of a portal's section into dynamic results.
- **Sector:** higher education.
- **Country of origin:** Spain.
- **Data sources:** open data from different national universities.
- **Number of universities involved:** six (until July 2022).
- **Website:** www.universidata.es

UniversiDATALab was conceived by the team that created UniversiDATA, the open data portal for Spanish universities. The portal went live in December 2020 with the goal of making it easier for universities in Spain to publish open data and to support its reuse. As a public–private collaboration initiated by three public universities and the company DIMETRICAL, UniversiDATA has recently expanded to include six academic institutions. These not only share one single portal, but also use the same standardisation, harmonisation and publishing schedule of the data produced. UniversiDATA features a series of sections, such as the 'Laboratory' section, which is the seed of UniversiDATALab and the main channel through which the UniversiDATA team has **showcased the power of open data for higher education**.

What services does UniversiDATALab offer?

As a follow-up to one of the sections of UniversiDATA, UniversiDATALab serves as an **interactive repository of analytical applications of the open data published on the portal**. UniversiDATALab is currently in a closed alpha state, only accessible by members of the project, and will only go live at the end of 2022. The UniversiDATA team is still in discussion with stakeholders to identify the best research content to make available. Yet, a crucial part of the repository will certainly be expanded with dynamic versions of the research so far conducted within the scope of UniversiDATA, among which include:

- an analysis of intercity commuting;
- a study on 'teaching in breeding' (i.e. how many professors have been students at the same university);
- a retirement analysis to predict how many retirements will happen in a specific year and in which university departments;
- a study on gender differences in enrolment and course performance.

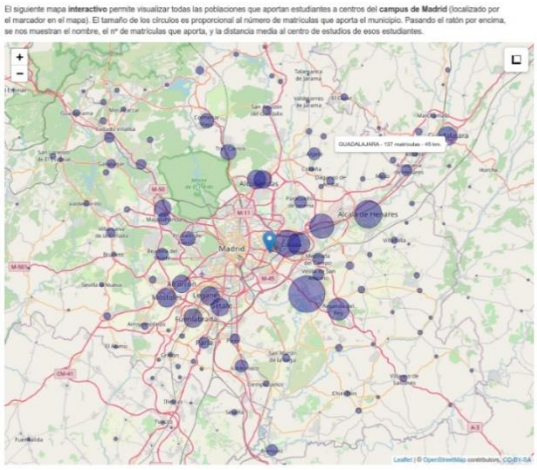
Further examples of open data applications might also draw from the outcomes of the first UniversiDATA Datathon that was launched in July 2022.

All the **applications on the repository will be made interactive and open source**. By doing this, the UniversiDATALab wants to materialise the potential value of the data published by universities through analytical applications that provide useful insights into the higher education community and serve as didactical tools of what can be achieved through open data. By leveraging the common semantics,

formats, granularity and schedule, the analysis implemented in these applications will also be instantly available to data published by every new university that adheres to the project.

UniversiDATALab aims at being as a source of inspiration and guide for a varied audience, including individual students using the data to make informed decisions on their academic path, teachers and researchers using the data as an academic resource, companies taking business decisions and government bodies relying on this data for monitoring and policymaking goals.

Figure 13. Example of an interactive map available on UniversiDATALab, depicting the hometown, the number of registered students and the distance to the campus for students who need to commute



What data does UniversiDATALab use?

As a spin-off of UniversiDATA, the Lab and the analyses published within it **will initially be based on the open data from UniversiDATA only**. However, additional open sources may be considered in the future. The data published on UniversiDATA is completely processed using the R Project for Statistical Computing, from the download of the data to its presentation via dynamic tools and graphs. UniversiDATA datasets are published at the finest level of detail, meaning that very large and detailed datasets are analysed and cleaned before this processing happens.

To enable this process to go smoothly, the UniversiDATA team has an annual schedule, defining when universities should provide them with the standardised data sources. Moreover, the team also has a technical guide on how universities should prepare these standard data sources so that the process of anonymisation and enrichment of data can be automatised and reused.

How does UniversiDATALab create an impact?

Since UniversiDATALab is not live yet, its impact is still difficult to predict. Nonetheless, the team plans to keep track of the number of visits and to introduce a login-free feedback mechanism, just like the existing one on the UniversiDATA portal. In this sense, the positive impact and the potential of the project could already be seen from the numerous comments appreciating the open data applications, which eventually led to the idea of creating the ad hoc repository.

The greatest ambition of the UniversiDATA team is continuing to have an **impact on universities' production and use of open data** and showing its potential through different applications. More broadly, the team also aspires to provide a useful service to as many collectives as possible and to include in the project an **increasing number of Spanish universities**. Several academic institutions have already asked for information and a demo of UniversiDATALab: including them in the initiative would **also improve the value, breadth and accuracy of the analyses**.

VisImE-360: the Italy-based app using open data to facilitate healthcare planning for visual impairment

VisImE-360 in a nutshell.

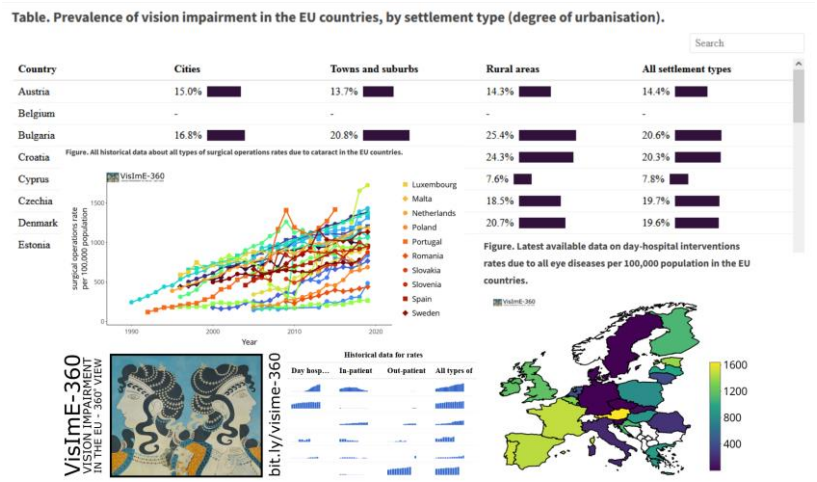
- **Service:** VisImE-360 explains in a single information space Eurostat's data on visual impairment, helping to allocate resources for medical aid.
- **Sector:** health.
- **Country of origin:** Italy.
- **Data sources:** open data from Eurostat and research studies.
- **Number of employees:** 1.
- **Website:** vision.scientific-tools.org

VisImE-360 was created by the epidemiologist and data scientist Boris Bikbov, driven by the idea to collect in a single information space major data on a single health condition. He used raw open Eurostat data to produce easy-to-percept text description, visualisations and tables. The 2021 EU Datathon competition provided the right opportunity for Boris to develop a web application that explains and visualises data on visual impairment in Europe. In this way, **the application ultimately helps decision-makers to allocate resources** for helping people with a visual impairment and **facilitate provision of the most appropriate medical aid**.

What services does VisImE-360 offer?

The goal of VisImE-360 is twofold. On the one hand, it **raises awareness about visual impairment**, which is one of the most prevalent health conditions worldwide, affecting an increasing number of people. On the other hand, it **supports a wide range of stakeholders** – including policymakers, patient organisations and people with a visual impairment, members of the media and social service workers – **to take action and improve access to prevention and treatment**.

Figure 14. Selected VisImE-360 visualisations



To do so, the application presents the user with two chapters. The first chapter shows the **number of people that suffer from a vision impairment** across Member States and according to the degree of urbanisation (i.e. people living in cities, towns or rural areas) and education (i.e. people with primary, secondary and tertiary education). This is done through different tables, bar charts and interactive maps. The second chapter focuses on **healthcare resources and utilisation for vision impairments**, including the latest metrics on available ophthalmic surgeons in Europe, up-to-date statistics on surgical operations for cataract treatment across Member States and data on day-hospital interventions and hospital discharges for all eye conditions. Dynamic line charts and tables help the user to draw comparisons between various countries in terms of prevalence and treatment patterns and use them to plan public health initiatives.

To make all this information more accessible to all – including individuals with decreased visual perception, colour vision deficiency or colour blindness – VisImE-360 offers the possibility to select **colour schemes and font features that better suit the user's needs** and allow them to have **a more comfortable experience**. Written and visual guidelines on how to explore the application and its features are also available.

What data does VisImE-360 use?

VisImE-360 entirely relies on **Eurostat data, scientific research studies and other open data sources**. To prepare the data and use it effectively for the application, Boris underwent a two-step process using the open-source R computing environment.

Firstly, he **searched through the datasets of Eurostat**. This was quite challenging because the initial search indicated almost 300 datasets with topics related to vision impairment and disability. However, only some of them contained information of interest. While some datasets could be filtered based on the complete description, Boris has to manually review about 150 datasets and their dictionaries to find those really containing data on vision impairment. The application was ultimately constructed on eight datasets.

The second part of the process entailed the **preparation of all visualisations, tables and figures to better reflect the richness of the open data** retrieved from Eurostat. The major findings are also accompanied by a short text description partially generated by a data processing algorithm. Importantly, Boris made all information processing and visualisations by using reproducible data science pipelines that facilitate their update upon new Eurostat releases in the following years. Through ad hoc accessibility features and different kinds of graphs, VisImE-360 offers easy-to-percept statistics for everyone.

How does VisImE-360 create an impact?

VisImE-360, which was first released in November 2021, is still up and running and has a development plan. An update of the datasets used is planned for this coming autumn and Boris would like to work on receiving financial support to refine the application and increase user engagement. At the moment, no precise analytics are available, but the app newsletter has 15 interested subscribers. At the same time, while no particular feedback was provided via the contact form of the app, VisImE-360 seems to have awakened a strong interest among the jury of the 2021 EU Datathon and its participants.

All in all, the **great benefit of VisImE-360 for visually impaired citizens and the larger health and policymaking community is certainly clear**, the application will be developed further and its impact is still to be assessed.

OpenActive: getting more people active with open data

OpenActive in a nutshell.

- **Service:** OpenActive is a community-led initiative coordinated by the Open Data Institute. Our aim is to make data available so it is easy to find and book a sporting or physical activity.
- **Sector:** sport, health.
- **Country of origin:** United Kingdom.
- **Data sources:** data from national leisure centres and sporting bodies.
- **Number of employees:** 2 to 10.
- **Website:** www.openactive.io

OpenActive is a community-led initiative working across the physical activity sector and backed by National Lottery and grant funding from Sport England and the Department for Digital, Culture, Media and Sport. Now entering its fifth phase, the initiative is coordinated and hosted by the Open Data Institute. OpenActive's journey started in 2015 in response to the fact that more than 40 % of the English population struggled to have an active lifestyle, a significant barrier being finding physical activities in their local area. OpenActive helps to address this problem by **helping its community members to publish standardised open data on physical activities and to promote innovation in making it easier to find and book these opportunities online**.

Figure 15. OpenActive logo

What services does OpenActive offer?

To achieve its main goal of encouraging people to be more active and enabling easier access to sport activities, **OpenActive developed a standard format to publish and use data about sport and physical activity opportunities**. Activity providers have two ways of promoting their activities to people, either by inputting this information directly into their own applications and websites (i.e. as data providers) or by relying on software tools that distribute it to further gathering applications and websites (i.e. via system providers). However, activity, data and system providers often use different formats for publishing their information, which can make it harder for people to find sport activities to book and attend. The open data standards developed by OpenActive help overcome this problem, benefitting both activity providers and the wider community.

While being born as a standards initiative, OpenActive is increasingly focusing on **supporting organisations to better understand the power of open data**. This includes helping activity providers to use the data standards, and building skills and capabilities across the sector to innovate with data to address the specific problems faced by people wanting to get active.

What data does OpenActive use?

The main OpenActive data providers are leisure centres (civic gyms) and sporting bodies, who tend to already have software in place to track activities and members. In the United Kingdom, this market is dominated by two or three software providers that specialise in leisure management and membership systems. Because of this, the main focus of phases one to four of OpenActive, starting in 2015, was to work with these providers to implement the OpenActive standards and support them to open up their data on sport and physical activity opportunities.

Another critical data source are the smaller organisations with no management systems, or sometimes even individuals (e.g. a yoga instructor). A member of the OpenActive community created a simple data entry application, called Open Sessions, that supports such providers in sharing information on activities offered, to allow these to be more widely advertised as open data. Other organisations such as the national governing bodies for specific sports are now publishing their data conforming with OpenActive standards – data that would have otherwise remained in silos. Opening up all this data allows users to fully reflect the varied choice of opportunities available in a local area.

While OpenActive focuses on developing and promoting the standards, **commercial partners are supporting implementation, cleansing data and providing middleware services**. OpenActive also provides information on whether data is conformant and some tools to allow organisations to undertake the data cleaning themselves.

How does OpenActive create impact?

In the last 7 years, OpenActive has had a tangible impact on the British physical activity sector, with **data on more than 200 000 opportunities for physical activity being published every month from 1 200 sites and 70 organisations**. This is mainly in England – especially urban centres like London – but also in Wales and Scotland. While only around 5 % of the supported organisations provide bookable activities at this stage, data is regularly being published and organisations are being empowered by OpenActive's open data approach.

In the past few years, the focus in the sector has shifted from organised sport in a conventional sense to physical activity and movement more broadly. This is reflected in Sport England's 10-year strategy – Uniting the Movement – and in the wider objectives of the fifth phase of OpenActive. OpenActive strives to **have a greater impact on under-represented demographics, looking beyond the big gyms at other use cases for OpenActive data**. In particular, looking across the sector to identify the potential for innovation that addresses specific user problems, for example, where there is inequality in access or experience, such as that faced by people with disabilities, or ensuring positive experiences for children and young people as the foundation for a long and healthy life.

In this broader context, Tim Corby and Howard Askew, from the Open Data Institute's OpenActive team, also envisage expanding the community network and encouraging more people to explore OpenActive data in relation to a broad range of challenges, for example, in supporting active travel or promoting social prescribing through links with the Open Referral UK standard.

Tangible data: letting citizens experience data by making datasets physical instead of digital

Tangible Data in a nutshell.

- **Service:** Tangible Data transforms data from its digital context to a physical context by creating data sculptures in the public space. These data sculptures help people that lack certain digital skills to experience the data.
- **Sector:** culture.
- **Country of origin:** Spain.
- **Data sources:** open data from international bodies (e.g. the National Aeronautics and Space Administration (NASA), the World Bank) and further platforms.
- **Number of employees:** 2 (no FTE).
- **Website:** www.tangibledata.xyz

Tangible Data strives to address the issue of lack of knowledge, misinformation, data accessibility and literacy in relation to sustainability challenges. According to Tangible Data, we are failing to get the best out of the data available to us. By making data touchable and physical, it is trying to bridge the data divide between groups of people without digital and data skills.

What services does Tangible Data offer?

Tangible Data was founded by Antonio Moneo and is one of the participants in the 2022 EU Datathon competition. Antonio observes that global warming is being denied by large groups of people and conspiracy theories are also coming back into vogue. Apparently, as a society we fail in informing certain citizens about the available data about these topics.

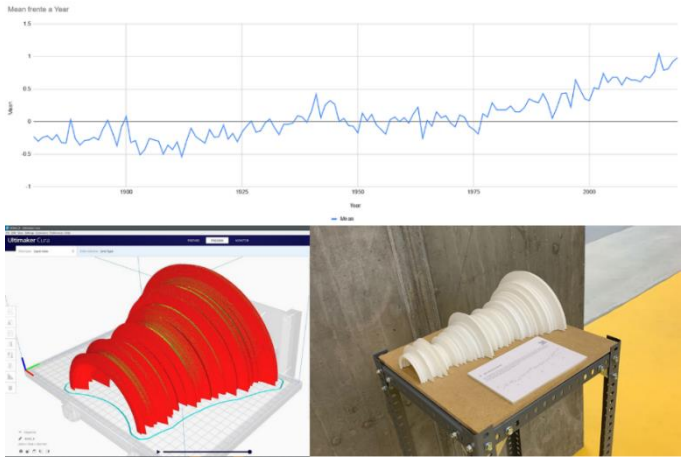
Data has the potential to transform our world by providing evidence to understand the needs and the impact of sustainable policies. But despite important advances in visualisation techniques, digital barriers significantly limit the impact of data. Instead of making data as easy to understand as possible, many organisations are busy building complex interactive dashboards. **This project aims for a new approach, using three-dimensional, physical representations of datasets to reach non-digital audiences in non-digital contexts**. In other words, Tangible Data tries to move data from the digital to the physical context. By making data sculptures, people can experience the data. These sculptures can then be placed in front of government buildings to inform citizens about certain subjects.

An interesting example is a prototype of the climate change tunnel that Tangible Data has already made (see Figure 17). The data in the figure shows the increase in average temperature on earth since 1880, leveraging data provided by NASA. The greater the warming, the larger and more spacious the tunnel becomes. The prototype of the tunnel is just 40 centimetres high, but this tunnel could easily be made a few metres high. If the tunnel is sufficiently large, then people can walk through the tunnel and experience the scale of global warming.

Other sculptures deal with the worldwide reduction of poverty and trust in public institutions. The aim is to help visitors familiarise themselves with relevant data about sustainable development challenges.

Tangible Data makes the sculptures using 3D printing and laser cutting techniques, which are highly scalable. This also makes it possible to produce the sculptures locally through involving the local makers community.

Figure 16. From data about climate change to a digital visualisation to a 3D prototype



What data does Tangible Data use?

Most of the sculptures use open data, which is easy to find and has a licence for reuse. For example, a dataset from NASA was used to illustrate climate change, one from the World Bank was used to describe the evolution of extreme poverty and data from Our World in Data was used to describe changes in trust in public institutions. In some cases, company data is used when a certain company wants to have a special visualisation for itself.

Tangible Data only applies the necessary transformations to the raw data in order to produce the 3D model and, in some cases, soften the edges of the sculptures and make them more touchable for audiences with visual impairments. The real challenge lies in creating a 3D model that visualises the impact.

How does Tangible Data create an impact?

Raising awareness about sustainability challenges is the first and most natural impact of the project. Data sculptures are designed to increase the visibility of certain topics and trusted data sources that provide accurate information about that specific topic.

A second impact is about empowering communities to leverage the use of data and technology. For example, Tangible Data is currently working with: groups of people with visual impairments to promote innovative approaches to data; teachers/professors (from nursery through to university) to design innovative teaching materials; and artists of different kinds to discover new intersections between art, technology and data.

Thirdly, Tangible Data aims to engage citizens in the solution of sustainable development challenges. Every data sculpture is equipped with a QR code that directs audiences to a website, where visitors will be able to purchase replicas of the data sculptures and contribute to the cause represented by the data sculpture.

In the coming years, **Tangible Data wants to make large data sculptures that can be exhibited in public places**. Tangible Data is also looking for funding to continue its work. To this end, it also works together with organisations interested in developing a data sculpture.

Hale & Hearty: the knowledge-based application improving the health and well-being of Irish citizens

Hale & Hearty in a nutshell.

- **Service:** Hale & Hearty is a knowledge base created by the Irish government to make health and well-being information more accessible and incentivise citizens to a healthier lifestyle.
- **Sector:** health, well-being.
- **Country of origin:** Ireland.
- **Data sources:** open data from data.europa.eu, local authorities, sport organisations, etc.
- **Number of employees:** 2–3.
- **Website:** haleandhearty.ie

Hale & Hearty is an EU-funded project managed by the Open Data Unit in the Department of Public Expenditure and Reform of the Irish government and supported by the Central Statistics Office, the Department of Health, the Fingal County Council, Derilinx and Ireland's Data Cloud Cluster. It aims at **making health and well-being data open for analysis, insights and action**, through a knowledge base that links data from local authorities, central statistics offices and healthcare providers. To further incentivise people to learn from this public information and lead healthier lifestyles, the project also involves the creation of a mobile application.

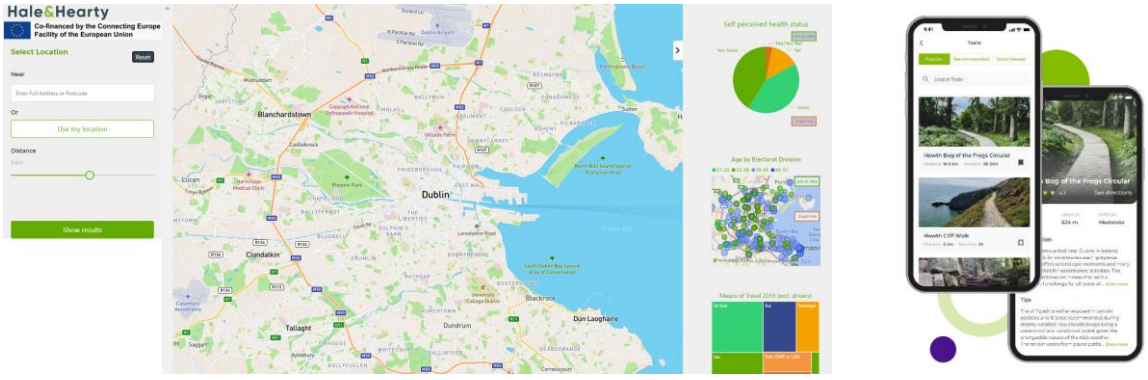
What services does Hale & Hearty offer?

In August 2019, the Open Data Governance Board of the Irish government proposed to create a knowledge base to make health and well-being information openly available. The idea was to allow the general public to use this information to improve their personal health and sport activity, to allow data professionals to better contextualise health statistics and to enable health professionals to make use of this data for new innovative services.

Financed through the EU Connecting Europe Facility grant and supported by six partner organisations, the Hale & Hearty knowledge base offers an extension of the Irish Open Data Portal (data.gov.ie) by providing up-to-date API access to high-value datasets and statistics related to the health and well-being sector, including Central Statistics Office (CSO) demographic and mobility data, local governmental data, open systems interconnection data on the location of sports and amenity facilities, network traffic analysis (NTA) data and health research data.

Hale & Hearty also offers a mobile application, to give the general data user further means by which to access the data, along with the choice to add their own anonymised data to track their activity and give mobile access to health and well-being facilities information. This allows users to access the data of the knowledge base to find local activity trails more easily in their area and connect their wearable devices to track their sport activities. The application also works as an incentive and gamification engine insofar as it allows users to earn points for each step they take, win bonuses, redeem vouchers and compete with other users.

Figure 17. Hale & Hearty knowledge base and mobile application



What data does Hale & Hearty use?

The data used for the Hale & Hearty knowledge base and mobile application is completely open. In fact, this data is published on data.gov.ie if it is not already there. Different data sources are consulted to retrieve this data: data.europa.eu, sports organisations, the Health Indicators Service of the Irish Health Service Executive and local authorities. All the data is published under the CC BY 4.0 licence.

In general, data processing is carried out by the data providers. In the case of the CSO data, entries were added to the data catalogue using the CSO API. Other data from data.gov.ie was flagged for inclusion using the group feature and a harvest process was set up to publish the Hale & Hearty group datasets on the data catalogue. This means that any updates to these datasets are reflected in the data catalogue and knowledge base when the harvester is run.

Once a month, an audit is done of any new datasets published. Once relevant datasets are identified, they are added to the Hale & Hearty group on data.gov.ie and are published on the Hale & Hearty data catalogue by the harvester.

To present the Hale & Hearty user with a menu of data that they can filter to select data of interest, the data was also categorised.

How does Hale & Hearty create an impact?

The Hale & Hearty team regularly monitors the number of hits the knowledge base receives. In January 2022, they ran a survey to gather feedback on the website and on potential developments of the mobile application. The knowledge base splash page also includes a request to participate in a survey on its usefulness and relevance. The team also held workshops for healthcare professionals to assess how relevant and useful Hale & Hearty is to their requirements. Despite this, it is difficult to objectively monitor the impact of the Hale & Hearty action on Ireland's health.

Looking ahead, the sustainability of Hale & Hearty requires substantial investment of expertise and resources. The knowledge base will continue to be available to the public and data will be added and updated on it. The Hale & Hearty mobile application requires further development to be sustained and made widely available. It needs to become a tool for a large-scale health service programme of activity. This element of the project requires a sponsor to actively manage it as it is outside the current capacity of the Open Data Unit.

EU Twinnings: exploring similar regions across Europe with open data

EU Twinnings in a nutshell.

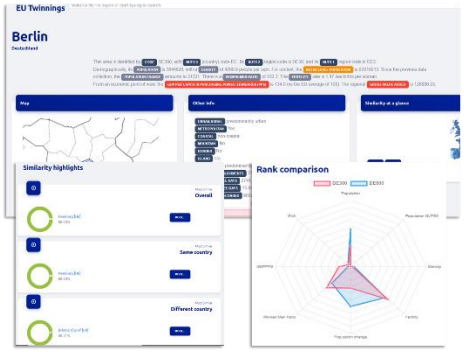
- **Service:** EU Twinnings uses open data from Eurostat to make statistics accessible to a wider audience and show similarities across EU regions.
- **Sector:** society, European integration.
- **Country of origin:** United Kingdom.
- **Data sources:** open data from Eurostat and data.europa.eu.
- **Number of employees:** 1.
- **Website:** data-europa-eu.eu-twinnings.site

EU Twinnings is the visually catchy application that the British-Italian data scientist Giuseppe Sollazzo presented at the EU Datathon competition in 2020. The application is based on an idea that Giuseppe firstly developed when reading an academic paper that calculated the degree of similarity between literature pieces: but what if the same concept could be applied to European countries? EU Twinnings realises this **question by allowing people to explore and visualise Eurostat's statistics on specific regions and municipalities and compare their similarity** with other European territories.

What services does EU Twinnings offer?

EU Twinnings web application is not a commercial, but an exploratory idea through which statistics and open data are made more understandable and the similar characteristics of European countries are highlighted. To do so, the app uses a formula (i.e. the cosine similarity, which is a measure of similarity between two sequences of numbers), which produces a percentage of similarity for European regions. In practice, the user can select a specific area or municipality in Europe and visualise the extent to which that particular area is similar to others in Europe, according to socioeconomic parameters such as population, gender balance and gross domestic product. Below a first overview dashboard on the selected area, the app provides 'similarity highlights'. By clicking on 'more', a rank comparison in the form of a spider chart on a specific similarity highlight and across parameters can be visualised.

Figure 18. EU Twinnings' interface



What data does EU Twinnings use?

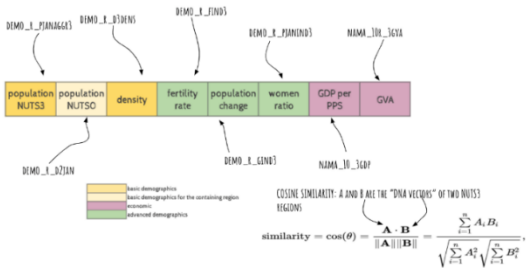
EU Twinnings is completely based on **open data from Eurostat**. The datasets used are considered both from a **NUTS 2 (nomenclature of territorial units for statistics) and NUTS 3 level** (i.e. they refer to the

current NUTS 2021 classification valid from 1 January 2021 and listing respectively 242 regions at NUTS 2 and 1 166 regions at NUTS 3 level). The datasets retrieved from Eurostat relate to socioeconomic parameters such as population density, fertility, gender, and gross domestic product. Sometimes, however, the same coverage of data is not available for every region and finding parameters of similarity may become complicated. Therefore, while Eurostat provides access to a variety of different data and the open licence allows this data to be used in a very easy way, Giuseppe is considering also relying on other open data, for example, from local census.

For both NUTS 2 and NUTS 3 regions, Giuseppe prepares a list of 10 **demographic and socioeconomic parameters** and a list of regions for which those parameters are available. He uses the EurostatJSON API to download data from Eurostat and undergoes cleaning and unit testing. Following this, he calculates the similarity between two regions by applying the similarity formula selected by him, which produces a percentage extent to which those two regions are similar based on the definition used for similarity.

EU Twinnings' **similarity formula** is automatically reviewed once a year, as the data for NUTS 2 and NUTS 3 regions is periodically updated in Eurostat. What Giuseppe still finds difficult to fully automatically adapt are updates of the classifications of the borders' areas and changes to tables in terms of terminology.

Figure 19. Similarity formula used by EU Twinnings



How does EU Twinnings create an impact?

Given the non-business driven focus of EU Twinnings' idea and in light of privacy reasons, the app does not identify its daily and monthly users through specific analytics and it is therefore hard to quantify its impact on a wider audience. On the other hand, the impact of EU Twinnings is qualitatively proved by the strong interest awakened by the application during the 2020 EU Datathon. Following the event, Giuseppe received several suggestions, for example, to use EU Twinnings within the framework of the Erasmus+ programme as a tool for mobility students. The impact and scalability of the application is also demonstrated by the new 3-year **collaboration between EU Twinnings and data.europa.eu**, the European data portal managed by the Publications Office.

In this context, Giuseppe is working on some **new EU Twinnings features, including a feedback button** through which users can share their input and contribute to improving the application. In terms of improvements, Giuseppe has also just concluded a **revamp of the design of the application** and is considering new definitions of similarity to keep the formula of EU Twinnings as inclusive and updated as possible. Similarly, he is not excluding the idea **of integrating data on tourism and criminality** and possibly buying this data from non-open repositories. In fact, sources of open data – especially for climate or academic data – do not generally have a level of granularity useful for an app like EU Twinnings. Hence, it is necessary to understand at which level data can be retrieved and if this data is worth buying to better calculate the similarity of European regions.

Open Food Facts: informing the consumer about the nutritional and ecological effects of their food

Open Food Facts in a nutshell.

- **Service:** Open Food Facts creates easy-to-understand information about the nutritional value and the environmental impact of food, together with a large food product database containing over 2.5 million products.
- **Sector:** food, health.
- **Country of origin:** France.
- **Data sources:** open data from food producers and further national and European sources.
- **Number of employees:** 7.
- **Website:** fr-en.openfoodfacts.org

The food we buy can have negative effects on our health and on the environment and we, consumers and manufacturers alike, often lack the information to make the right decisions. Open Food Facts is a **large open data base for food products where users can find information about the ingredients of certain products**. Using other open data sources, Open Food Facts creates easily understandable scores for nutritional value, environmental impact and the extent to which the product contains processed foods.

What services does Open Food Facts use?

For Open Food Facts, the goal is to make food data a common good that can be used to challenge some of the large problems that developed countries are currently facing, such as the obesity crisis and global warming.

For the first 7 years, all the work that contributed towards Open Food Facts was done pro bono, but over the last few years it has been able to receive grants from the French National Health Agency and other institutions (e.g. under the Commission's next generation internet initiative and private foundations). Furthermore, it recently received the Google Impact Challenge on Climate, which made it possible to further expand the organisation.

Currently, Open Food Facts is a **database containing 800 000 products in France and 2.5 million products worldwide**. On the website, users can find most of the information that can also be found on the label of the product if they were to buy it in the supermarket. Information such as the ingredients the product contains and whether the packaging can be recycled. Moreover, the website shows in which supermarkets and in which countries the product can be bought.

It is not an easy task to look at the ingredients and understand whether a certain product is good for you or not. That is why Open Food Facts also adds the nutri-score label on their website. The nutri-score is an attempt to simplify the nutritional value of food products on a scale from A to E and can be calculated based on the ingredients in the product. Over the last couple of years, **the nutri-score has become widely adopted across France** and food producers even reach out to Open Food Facts to discuss their nutri-scores.

Open Food Facts is in the middle of rolling out a so-called eco-score together with a number of non-profit organisations and start-ups. The eco-score is an environmental score from A to E that makes it

easy to compare the impact of food products on the environment. The eco-score is based on a life cycle analysis of the product, combined with data that can be found on the label, such as the mode of production, the origin of ingredients, whether the product uses packaging that can be recycled and whether the product has a negative impact on endangered species (for instance by using palm oil).

Lastly, Open Food Facts also provides a nova-score, which is another simplified measure constructed by food scientists to show the extent of processing involved in creating the product, given that ultra-processed foods have negative health effects.

What data does Open Food Facts use?

The data on Open Food Facts is collected either by the users of the platform or by food producers that want to have the data on their products available on the website. First, the users entered all of the information on a label into Open Food Facts, or they shared a picture of a label that was then manually uploaded by the Open Food Facts team. Now they have **an algorithm that automatically reads the information on the label**, sparing the team lots of time. Producers can add their data via the producer's platform. Smaller organisations can enter their data manually or via Excel, whereas big food producers (such as Nestlé and Ferrero) can use automated means to upload the data on their new products.

Open Food Facts is not just an open database. It also uses other open data sources to complement their analysis. They use European sanitary numbers on labels to trace in which factory a product is made and data from the European Food Agency to find the risks of overexposure to certain additives. Importantly, for the new eco-score the life cycle analysis data is from the open Agribalyse environmental database designed by the French Agency for Ecological Transition (ADEME) and the French National Research Institute for Agriculture, Food and the Environment (INRAE). Without these open data sources, Open Food Facts would not be able to provide their nutri-score, eco-score and nova-score.

How does Open Food Facts create an impact?

The biggest impact Open Food Facts has had is through the adoption of the nutri-score in France. What started as a niche rating system developed by food scientists is now on almost all products in French supermarkets. Food manufacturers actively look for their advice on how to make their products more nutritional and eco-friendlier. Their goal is to take the nutri-score global and to create a similar impact with their eco-score. Another addition could be to take a better look at the prices of food: which parties in the supply chain earn what?

It should be noted that everything Open Food Facts produces is open: **its data is open and downloadable via an API and its algorithms and source code are open as well**. Scientists often use the nutri-score data for research on overweight, and so do many other organisations. Currently over 150 applications use the data from Open Food Facts for several purposes, such as helping pregnant persons, people with disabilities or people with certain allergies to choose products that suit them. The large food and cosmetic scan application Yuka also started off with Open Food Facts data and still shares its data with them.

Integreat: the German platform helping municipalities to integrate migrants and refugees

Integreat in a nutshell.

- **Service:** Integreat is a digital platform that provides all relevant information in several languages at the municipal level to newly arrived migrants and refugees.
- **Sector:** society, migration.
- **Country of origin:** Germany.
- **Data sources:** open data from national municipalities.
- **Number of employees:** 27–35.
- **Website:** integreat-app.de

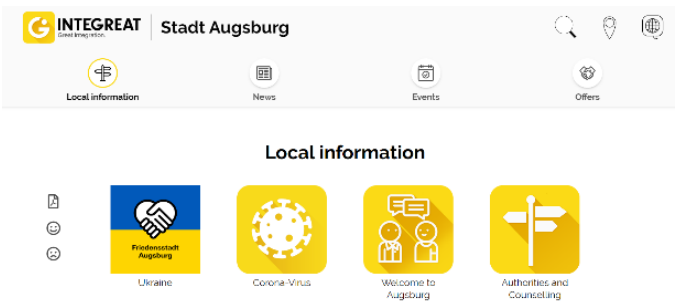
Integreat started in 2015 in Augsburg under the name 'Refguide+' to help the Bavarian city cope with the wave of incoming refugees and the information and linguistic gap that came in consequence. Soon after that, further municipalities in Germany reached out to the Integreat team to implement the solution. While the report was being written, in July 2022, Integreat was helping **90 municipalities make information for the integration of people with a migrant background publicly available** in different languages. Integreat therefore functions both as an integration process tool and as an incentive to further open key data at the municipal level.

What services does Integreat offer?

The idea behind Integreat is to gather all possible and useful information that will allow refugees and migrants in Germany to settle and integrate into the new society more easily. To do so, the solution offers refugees and migrants **public access to different data and information provided directly by the municipalities** where they reside, such as administrative processes, job opportunities, education and social service facilities, along with more recent COVID-19- or Ukraine-related information. Refugees and migrants can access this content in multiple languages either on the app, website or offline through ad hoc brochures.

All information on Integreat is provided by municipalities and experts in the area, who are also responsible for the maintenance of this data: no IT know-how is required to add and maintain information on Integreat as the content management system is easy to use. The Integreat team, on the other hand, helps municipalities with the first implementation of the solution and remains available for advisory and tech support on publishing the data. Moreover, the team works to create more **opportunities for exchange with other cities and districts and to foster the sharing of standardised information** among them, especially when it comes to legal information.

Figure 20. Integreat app for the municipality of Augsburg



What data does Integreat use?

Integreat is completely based on open data. The idea is to decrease information poverty by making data on housing, healthcare, education, family services, free time, work and more recent topics such as the COVID-19 pandemic or the Russian war of aggression against Ukraine available and easily accessible everywhere. This data is **common knowledge published on Integreat directly by municipalities**. The complete programme and source code of Integreat is in fact freely available and provided under an **open-source licence (MIT)**. Moreover, all content of the various municipalities is licensed under the Creative Commons (CC BY 4.0), meaning that new and existing partners can benefit from each other in terms of content and translations, and workload in creation and maintenance is minimised. There is no need for the Integreat team to manipulate the data. However, the team offers tech support and advisory services to municipalities on how to best publish the information and maintain the Integreat platform.

How does Integreat create impact?

Since its development, Integreat has been constantly expanding to new areas, reaching **90 partner municipalities across all of Germany** in July 2022: in Bavaria, every third municipality currently uses Integreat and the solution is also well spread across North Rhine-Westphalia and Hessen. Given the different political structure, however, Integreat has not yet reached many municipalities in the north of Germany and is so far not at all implemented in Thuringia and Mecklenburg-Vorpommern. While in the first 3 years since its development Integreat was supported by public funding, since 2018 municipalities are required to pay between **EUR 5 000 and EUR 7 000 a year to adopt the solution**. In 2021 the total revenue of Integreat amounted to around EUR 240 000.

Besides the good geographical spread and the encouraging economic development of the last few years, Integreat's positive impact on municipalities is certainly proven by the fact that, since the beginning, very few partners have cancelled their subscription to the platform (main reason for cancellation being personnel shortages). The Integreat team is very keen on maintaining a **good relationship with the municipalities** and organises **qualitative feedback talks** with them once a year. This has allowed Integreat to distinguish itself from competitors, offering not only technology that is second to none, but also becoming very close in the collaboration with municipalities.

Regarding Integreat's impact on its end users, namely refugee and migrant communities in Germany, the evaluation is more complex, as downloads are not sufficient to properly track the usage and the platform can also be accessed offline. Yet, the Integreat team is currently conducting a **randomised control trial study funded by J-PAL Europe to see how the use of Integreat is changing migrants' and refugees' ability to access services** in their municipality of residence and which information within Integreat is most useful for them. The first results should be published in September 2022.

Looking ahead, the Integreat team is planning several updates for the platform. Firstly, they would like to offer a **better automatic translation** to reduce translation costs for municipalities. Secondly, they aspire to make it **easier for Integreat to be integrated into the websites** of other local organisations. Thirdly, they are in the process of **piloting a chatbot** as an alternative way of finding information. Finally, they would like to offer migrants and refugees the **possibility to look for jobs** directly through Integreat and **scale the solution beyond Germany**. As a first step, they are considering implementing the solution in Greece.

Emergency Volunteer Application: Belgian app uses open data on automated external defibrillators to save lives of citizens when they have a cardiac arrest

EVapp in a nutshell.

- **Service:** EVapp sends volunteers with first aid training as quickly as possible to someone having a cardiac arrest. These volunteers often get there before the ambulance staff and thus save lives.
- **Sector:** health.
- **Country of origin:** Belgium.
- **Data sources:** open geodata and anonymised data from national emergency centres.
- **Number of employees:** 12.
- **Website:** www.evapp.org

The first 10 minutes after a cardiac arrest are the most important, and ambulances are not always able to arrive within 10 minutes. EVapp uses a **network of first responders to have someone on the spot as soon as possible to perform resuscitation**.

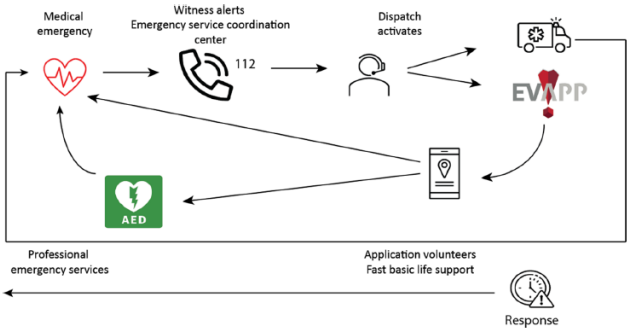
What services does EVapp offer?

EVapp was founded in 2015 by the company Prior-IT and EVapp NGO as a non-profit association to provide faster help to people having a cardiac arrest. Since 2017, the app has been live in the Hoogstraten region of Belgium. The creators are still waiting for Belgian legislation to be able to roll out the app to other parts of Belgium. Hoogstraten was chosen because it is one of the most difficult places in Belgium for ambulances to reach.

The app works as follows: as soon as someone has a suspected cardiac arrest, is drowning or is electrocuted and the Belgian emergency number 112 is called, the emergency centre forwards this message to the ambulance and to EVapp. The app then automatically alerts five citizens with first aid certificates who happen to be in the vicinity of the victim. **Two citizens receive the location of the victim and must reach them as soon as possible to perform cardiopulmonary resuscitation (CPR)**. They often reach the victim before the ambulance. Other volunteers receive the location of the nearest automated external defibrillator (AED) and must collect the AED as quickly as possible and take it to the victim.

In Belgium, anyone can obtain a first aid certificate. These certificates are valid for 5 years. As soon as the certificate is no longer valid, the app users are no longer asked to come to the location of the accident. When the certificate is obtained, the app's existence is brought to the attention of the certificate holders. Currently, 2 700 volunteers are connected to the app, 250 of which live in Hoogstraten.

Figure 21. Schematic representation of the EVapp system



What data does EVapp use?

EVapp makes use of open software as much as possible. For example, it uses OpenStreetMap to locate the victim and the first aiders. In addition, Prior-IT is working on mapping the locations of all AEDs in Belgium. Part of this data is open: they received a list of 8 700 AEDs and their locations from the government and additional data from several cities. Unfortunately, in some cities, almost half of all available AEDs were not part of the data provided by the government. AED manufacturers shared data on which organisations bought AEDs from them, leaving Prior-IT to fill the gaps. **Prior-IT contacted these organisations to confirm the presence of the AED.**

The AED data still needs to be enriched. For instance, AEDs in shops are only accessible when the shop is open. The opening hours could be found using web scraping, but in many other cases it was necessary to contact the owner of the AED to ask for additional information. The enriched AED location data collected is made available as open data.

Prior-IT also receives data from the emergency centres. Only the location of the victim is shared for privacy reasons.

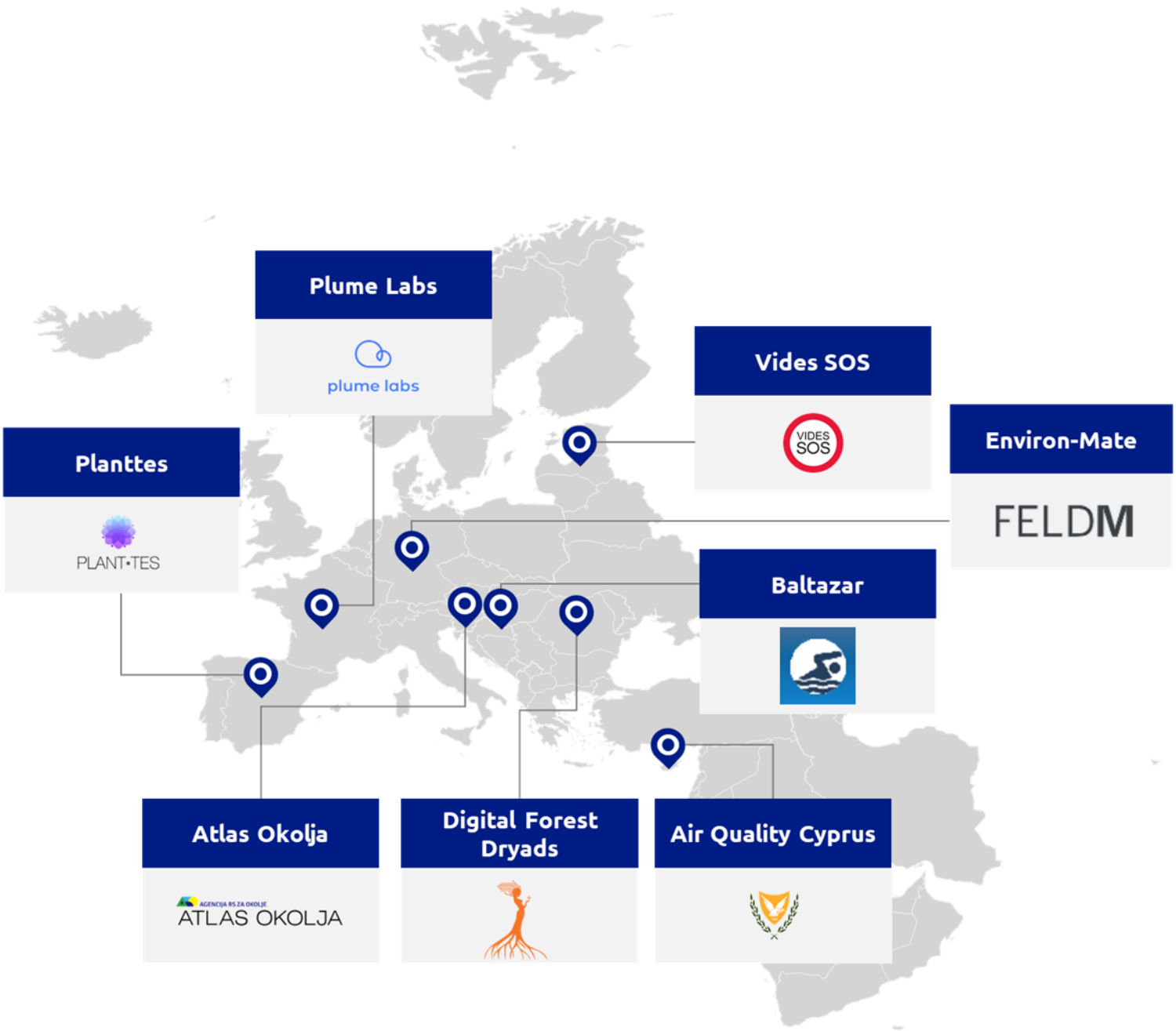
How does EVapp create an impact?

First estimates in Hoogstraten show that EVapp increases the chance of survival by 5 to 6 %. Moreover, **simulation of nation-wide EVapp implementation resulted in an additional yearly 910 quality adjusted life years (QALYs)** gained over the current baseline case scenario with a lower estimate of 632 and a best-case scenario of 3 204. Moreover, the cost per QALY would also decrease as a consequence of both the lower healthcare costs for patients with good neurological outcomes and the more efficient use of available resources ([5]). After each incident, EVapp collects feedback on the functioning of the app and the accuracy of the alarm. There is also emotional support available for volunteers in case the victim did not survive their cardiac arrest.

In the upcoming years, EVapp wants the app to become available nationwide. Furthermore, Prioir-IT is working on another app: the AED-hunter. This app should show all AEDs in Belgium on a map, so people can immediately see where the nearest available AED is on the app.

([5]) Vercammen, S. and Moens, E., 'Cost-effectiveness of a novel smartphone application to mobilize first responders after witnessed OHCA in Belgium', *Cost Effectiveness and Resource Allocation*, Vol. 18, No 1, 2020, pp. 1–11.

3.4. Environmental impact



Digital Forest Dryads: using open satellite imagery to protect forests from illegal deforestation

Digital Forest Dryads in a nutshell.

- **Service:** Digital Forest Dryads application aims to protect forests from illegal deforestation in Europe by combining aerial and multi-spectral satellite imagery.
- **Sector:** environment, forest.
- **Country of origin:** Romania.
- **Data sources:** open geodata and multi-spectrum data from several EU open sources.
- **Number of employees:** 5.
- **Website:** digital-dryads.eu

Digital Forest Dryads was developed by the Romanian team made up of Razvan Pistolea (chief technology officer), Andrei Mocanu, Sergiu Eftimie and Adnan Temur to participate in the 2020 edition of the EU Datathon. By relying on data from EU sources, such as Copernicus and Eurostat, and the Google Earth data catalogue, the team created an **application able to visualise an interactive map of deforestation and easily distinguish illegal activities from legal cutting**. Besides Romanian forests, the app also offers interactive maps of Albania, Belgium, Bulgaria, Croatia, France, Germany, Greece, Hungary, Italy and Spain.

What services does Digital Forest Dryads offer?

The goal of Digital Forest Dryads is to empower the forestry industry by helping both law enforcement authorities and ministries to **better analyse deforestation** and be able to take informed decisions against it.

In fact, the interactive map of Digital Forest Dryads allows users to **visualise which territories are covered by forest and in which areas the forest is disappearing**. More specifically, forest loss is represented by red areas on the map, while green zones refer to protected forests according to the United Nations Educational, Scientific and Cultural Organization (UNESCO) (i.e. areas where cutting is not legally permitted). By zooming on the map, it is also possible to see **yellow and red dots**: the yellow dots are used to indicate acts of deforestation that occurred legally, whereas red dots represent illegal actions.

Finally, Digital Forest Dryads offers the possibility to view deforestation in the selected country over time, visualise the state of protected forests and inspect legal deforestation.

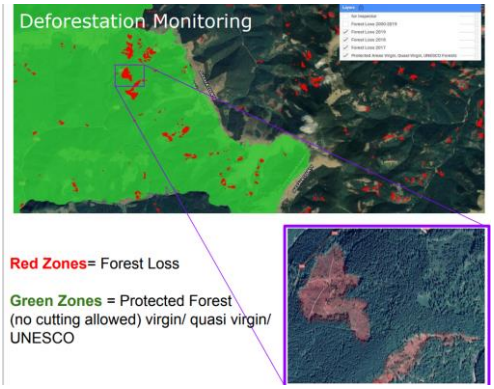
What data does Digital Forest Dryads use?

The team based itself on both open and non-open data. **Open satellite imagery** come from the Directorate-general for Defence Industry and Space, Copernicus, the European Organisation for the Exploitation of Meteorological Satellites, the Directorate-General for Health and Food Safety and Eurostat. Land images were also retrieved from the Google Earth data catalogue and Amazon Web Services.

Further data was obtained by the team through lobbying activities. For example, the information provided in the yellow dots concerning the legality of deforestation in a particular area – such as the Internet Protocol (IP), the cutting year, the entity responsible for the protection of the area and the

body that gave the permission to cut trees in the area – is not publicly open. The team was able to retrieve this data by writing emails to ministries, which then provided them with the needed geographic information system data.

Figure 22. Interactive map of Digital Forest Dryads



The data used by the web application, including the one coming from Amazon, is updated every 5 to 6 days and after cleaning was processed through machine techniques able to identify and classify trees.

Yet, the application is still based on **historical data from 2017, 2018 and 2019**. After the proof of concept and the first pilot, the team stopped processing data due to lack of resources – the project would require millions of euro (under **1 % of today's costs** that involve using traditional methods).

How does Digital Forest Dryads create an impact?

Since its launch, Digital Forest Dryads' impact can be observed on both a quantitative and qualitative level. In fact, the team regularly keeps track of the application's users through Google analytics. The monthly usage swings between **800 and 1 000 users per month**, with most of them utilising the desktop rather than the app version of Digital Forest Dryads.

The application is therefore well known among its target audience and attracted quite some interest from the Romanian government. Recently, the Romanian government decided to reuse the idea behind Digital Forest Dryads, which will be implemented through the financial resources provided by the EU Recovery and Resilience Facility.

Overall, the application has been **successful in raising awareness** about the deforestation problem in Romania (and in the rest of Europe) but is **lacking the support – especially economic – to be further updated**.

Despite this, following the release of Digital Forest Dryads, the team received **many collaboration requests and two other projects started**. A first project is 'wood watcher', an application through which the user is able to calculate the volume of a pile of wood from a simple picture. This is a very useful application for law enforcement authorities fighting illegal trafficking of wood. The application worked so well for Romania that the team released it to the entire world. Recently the team also presented the app during a conference organised by the International Criminal Police Organization (Interpol) on eastern European and Latin American practices and received the best practice in Europe award for their innovativeness.

Air Quality Cyprus: informs citizens about air pollution across the island

Air Quality Cyprus in a nutshell.

- **Service:** Air Quality Cyprus provides citizens with real-time information about several forms of air pollution. Users can find the data online or choose to be proactively informed about certain substances via the app on their smartphone.
- **Sector:** health, environment.
- **Country of origin:** Cyprus.
- **Data sources:** open data from national air pollution measurement stations.
- **Number of employees:** 5–10.
- **Website:** www.airquality.dli.mlsi.gov.cy

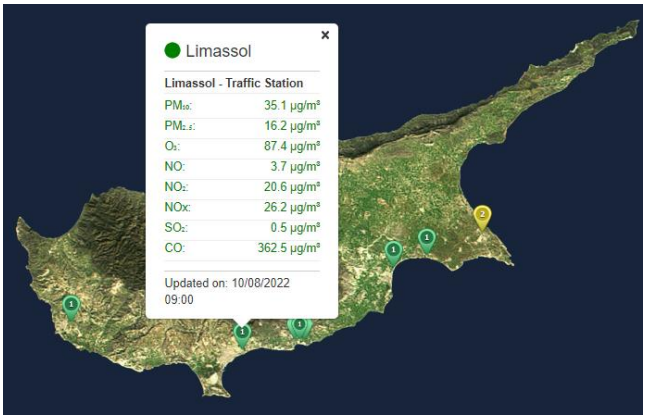
Air pollution can lead to respiratory problems, exacerbated allergies and have adverse neurological, reproductive and developmental effects. Vulnerable populations such as children, the elderly, pregnant women, those with heart or lung disease are especially at risk. In Cyprus, air pollution is caused by transport, industry and agriculture (to name a few), dust carried by the wind from three of the largest desert areas in the world (the Sahara, the Arabian Desert and the Syrian Desert). Air Quality Cyprus **helps users by providing real-time data about air pollution across the country** and information about the impact of certain forms of air pollution on their health.

What services does Air Quality Cyprus offer?

The Air Quality Cyprus website was launched 12 years ago to help Cypriots and people living in Cyprus to track air pollution in their area. The website is run by the Cypriot Department of Labour Inspection and the Ministry of Labour, Welfare and Social Insurance. In 2019, they also launched apps for Android and for iOS to make it easy for users to get real-time information on their smartphones.

The application and the website show a map of Cyprus and its nine air pollution measurement stations, which are spread across the country. Users can hover over the measurement stations and **find up-to-date pollutant concentrations** for eight polluters in the majority of the stations. Colour coding informs users about the risks of certain pollutant concentrations. The app also enables users to get notifications when pollutant concentrations reach a certain safety threshold in their area.

Figure 23. Information about air pollutants at the Limassol traffic measurement station



Air Quality Cyprus also provides information on the different forms and sources of air pollution, the adverse health effects for the different polluters, advice on how to improve air quality and an overview of existing legislation to prevent air pollution.

What data does Air Quality Cyprus use?

Air Quality Cyprus uses the **data from the measurement stations**. This data is also available on the Cypriot open data portal. The majority of the stations have instruments that measure the concentration of ozone, nitrogen oxide, nitrogen dioxide, nitrogen oxides, sulphur dioxide, carbon monoxide, benzene and particulate matters (PM10 and/or PM2.5), which are updated every 2 minutes in Air Quality Cyprus' software and twice per hour on the Air Quality Cyprus website and mobile application. Note that the data gets validated after instant publication on Air Quality Cyprus, otherwise **real-time information** provision would not be possible. Still, users gain a lot by having the unvalidated real-time information, since it is often a reliable estimator of the validated measurements.

How does Air Quality Cyprus create an impact?

Air Quality Cyprus **helps to prevent adverse health effects of air pollution for its inhabitants**. When air pollution reaches dangerous heights then Air Quality Cyprus informs the government, which subsequently alarms ministries (Ministry of Health, Ministry of Education, Sport and Youth, etc.), local governments, media and labour unions, who can then take immediate action. Precise estimations of the number of life years won because of the Air Quality Cyprus website and apps do not exist (and would be very hard to estimate), but both are used by many. The website has over **800 000 unique visitors per year**, the iOS app and the Android app have approximately 10 000 users each. Moreover, the website and the apps were updated recently to become more accessible for people with visual impairments.

In the near future, Air Quality Cyprus wants to create a forecasting tool to predict pollution concentrations. This would require them to also use (open) data from other organisations, such as meteorological institutes, to accurately account for wind force and direction, heath, rain and other relevant circumstances.

Vides SOS: the Latvian application to facilitate the reporting of environmental violations

Vides SOS in a nutshell.

- **Service:** Vides SOS is an application designed to alert the Latvian State Environmental Service of environmental hazards such as pollution, waste and litter in nature.
- **Sector:** environment.
- **Country of origin:** Latvia.
- **Data sources:** open data from OpenStreetMap.
- **Number of employees:** 6.
- **Website:** www.videssos.lv

Vides SOS is an open-data-driven application developed by the Latvian State Environmental Service to allow citizens to **report environmental infringement offences relating to air, water or soil and contribute to a cleaner environment**. The application relies on OpenStreetMap to visualise the areas where the infringements have happened. The app not only allows people's awareness to be raised about environmental protection, but it also incentivises collaboration between public institutions and citizens.

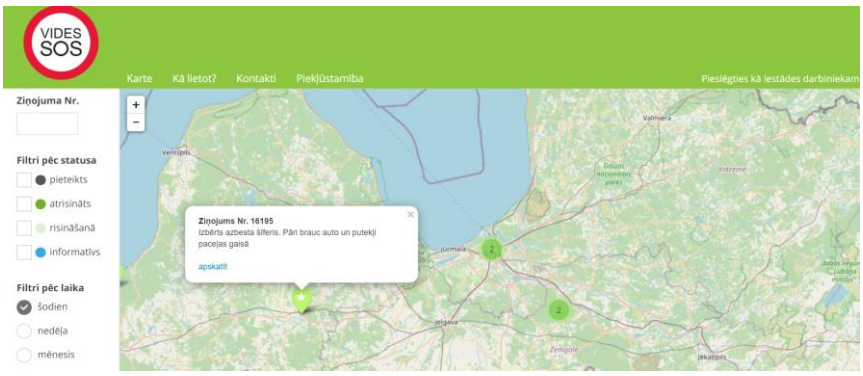
What services does Vides SOS offer?

Vides SOS is smartphone-friendly application that enables Latvian citizens to take a part in taking care of the environment by reporting air-related (odours, dust, smoke and radiation), water-related (activities of edge fishing, presence of chemicals, dead animals or oil) or soil-related (waste, manure and oil) infringements to the State Environmental Service.

The application therefore represents a **modern solution in the processing of environmental violation reports at centralised and coordinated level** and enhances public involvement and awareness. In fact, any citizen can simply download the application on their smartphone, take a picture of the infringement discovered and report it with a brief message either anonymously or with their account. In this way, the state or local institutions responsible for the area in which the infringement took place can contact the user for further information and act fast.

The infringements are then visualised on the map and can be filtered according to whether they have been taken care of and the date they took place.

Figure 24. Vides SOS's map of infringements



What data does Vides SOS use?

Vides SOS is completely based on **open data from OpenStreetMap**, a collaborative project to create a free editable geographic database of the world. OpenStreetMap® is open data, licensed under the Open Data Commons Open Database License by the OpenStreetMap Foundation.

To develop Vides SOS, the team responsible within the Latvian State Environmental Service did not need to process any open data, but just use the map from OpenStreetMap to get the coordinates of possible environmental hazards.

How does Vides SOS create an impact?

As Vides SOS allows information to be sent anonymously, the team does not monitor active users. Yet, it is possible to get a picture of the outreach of the application through the number of reports received. Since the beginning of 2022, the Latvian State Environmental Service has already received **2 345 reports of possible pollution**. Last year, in the same period, 1 876 reports were collected, which means there has been around a **25 % increase in reports** in 2022. From the received reports in 2022, more or less **35 % have already been solved, while 49 % have been sent to competent authorities**.

Originally the application provided a **'Clean up myself' option**, which encouraged the general public to get involved and clean up certain types of infringements themselves. However, since the option was not fully fleshed out and functioning, the team decided to remove it temporarily. With the creation of an ad hoc department responsible for managing the application since June 2020 and following the modernisation project completed in September 2021, the team plans on revisiting this issue and reintroducing this function in the future.

Overall, the application still has some technical difficulties that the team would like to address and has its limitations in assigning nature inspectors to watch every corner of the Latvian territory and ensure environmental protection everywhere. However, the application is showing great potential, with not only high levels of involvement from citizens and the State Environmental Service, but also from other governmental institutions. In fact, in August 2022 the application counted **56 other partner institutions**, all standing for more action and protection of the environment.

Planttes: using citizen science to inform people with allergies about which plants are in bloom

Planttes in a nutshell.

- **Service:** Planttes is a citizen science application that informs users about which plants are in bloom and whether this might affect people with pollen allergies.
- **Sector:** environment, health.
- **Country of origin:** Spain.
- **Data sources:** open data from the Point of Information on Aerobiology and open geodata.
- **Number of employees:** 2 (with the help of students).
- **Website:** www.planttes.com

Estimations of the number of people in Europe with pollen allergies are as high as 40 % in some studies. When certain plants are in bloom, people with allergies experience symptoms such as a runny nose, itchy and red eyes and sneezing. In severe cases the symptoms can even lead to extreme tiredness and the necessity to stay indoors during large parts of the flowering period. Planttes tries to **help people with pollen allergies by providing precise information on which plants are in bloom** and which types of pollen are in the air.

What services does Planttes offer?

Planttes is a citizen science project, meaning that it wants to involve citizens in the development of scientific projects. The idea stems from a group of design students and researchers on aerobiology. **The goal of the app is to better understand the relationship between the environment and allergic diseases, contributing to an improved quality of life for people who suffer from them**. At the same time, this would lead to a better understanding of how climate change affects people with pollen allergies and provide an opportunity for students to learn more about plants and their phenology.

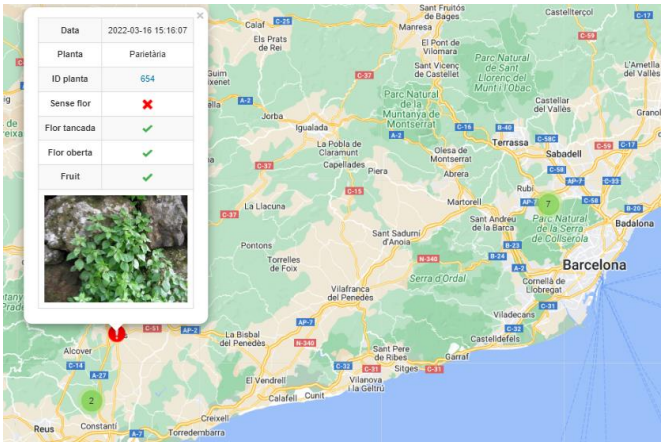
The Planttes team, together with an AI student, built an app that informs users about allergenic plants that might be in their surroundings. Planttes users can contribute to the elaboration of the map, which shows the phenological state (presence of closed flower, of open flower and/or fruit) of the plants in the vicinity that cause allergies. They can do so by indicating the place, selecting the plant, filling in the phenological state and uploading a picture to the app. Currently, the app predominantly has information from Catalonia, but it is useful in any area where the plants featured on the app grow.

The photos of the plants uploaded by the users are shown on a map (see screenshot). By accessing the pop up of each record, users can see the photo of the plant, together with the date of publication and the phenological data that determined the allergenic risk. The degree of risk is classified into three categories (low risk, increasing risk and maximum risk). By choosing specific filters for certain plants, users can create a personalised risk map.

What data does Planttes use?

Planttes uses data from two main sources: the data and pictures of flowers that people enter into the system and the open data on pollen from the Point of Information on Aerobiology.

Figure 25. Screenshot of the Planttes map with a warning for and a picture of an allergenic plant



The pollen data is collected via so-called pollen samplers placed on top of buildings and pollen analyses carried out under the light microscope by palynologists in the laboratory. **The data from the pollen samplers allow researchers to estimate which plants are currently in bloom**. In Point of Information on Aerobiology, the pollen data are then used to create a score between 0 and 4, where 0 means that people with pollen allergies should have no problems going outside and 4 means that their allergies will be intense that day. Other variables, such as the weather, are also included in this score.

The photos of plants that are currently in bloom and their location need to be validated before they are added to the map in the app. Currently, this is being done by two biology professors who check new entries regularly. The data is open after validation and can be accessed by any user via their website.

How does Planttes create an impact?

Planttes can create an impact in two ways: by improving the health of people with pollen allergies and by teaching people about plants. Their goal in the coming years is to regularly give presentations in schools, so that children can take photographs for the app and learn to recognise certain plants, and follow their development throughout the seasons, especially during the blooming periods.

Currently, the pollen information system provided by the Point of Information on Aerobiology is very useful and includes a large number of plants and fungi (around 700). However, the number of areas studied is insufficient to cover the entire country. Planttes therefore brings a new opportunity to citizens, showing them the reality of the blooming patterns of certain allergenic plants at ground level. **The app has been downloaded more than 1 000 times so far and the website is used by around 65 000 unique visitors annually**.

Atlas Okolja: everything you want to know about Slovenia in one map

Atlas Okolja in a nutshell.

- **Service:** Atlas Okolja (or environmental atlas) presents a map of Slovenia. The map combines information from a range of different sources, such as noise pollution, air pollution, earthquakes and Natura 2000 areas.
- **Sector:** environment, government.
- **Country of origin:** Slovenia.
- **Data sources:** open data from government bodies (e.g. cadastral, environmental).
- **Number of employees:** 2–3.
- **Website:** gis.arso.gov.si/atlasokolja

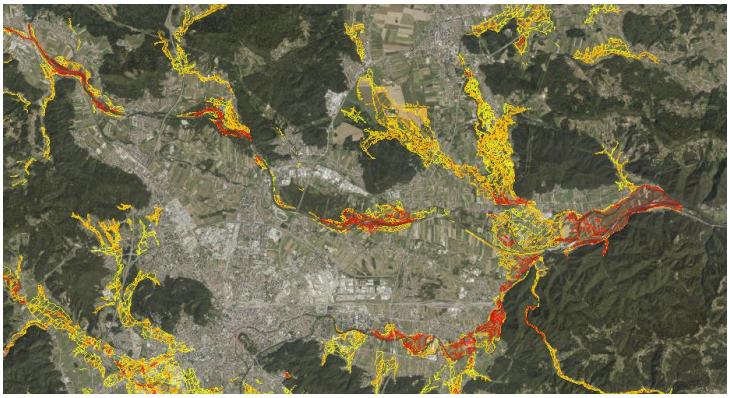
Atlas Okolja (environmental atlas in English) is a web tool for spatial data that was launched in 2008. The initiative is run by the Slovenian Environment Agency and **combines geo data from Slovenia with other data sources to add information to the map of Slovenia**. Examples of data sources that can be found in the Atlas are: mean temperature per region, average snow cover, noise pollution and maps for seismologic risks. The website helps citizens, city planners, insurance companies and researchers in making decisions, such as where to build new housing.

What services does Atlas Okolja offer?

Atlas Okolja helps Slovenian citizens find information about their country. The website provides a **satellite picture of Slovenia and provides users with a set of so-called layers that can be draped over the Slovenian map**. For example, the figure below shows the flood risks in and around Slovenia's capital Ljubljana.

The Atlas helps citizens, policymakers and entrepreneurs alike to make important decisions about land use in Slovenia. For instance, citizens that are considering buying land to build a new house. The Atlas data provides them with information on whether the land is flat enough to build on and whether the area surrounding their new house will be quiet and peaceful using noise pollution data. Moreover, city planners use data to see where restrictions for possible new housing exist, such as Natura 2000 areas and flood risks. Lastly, academic researchers use the data from Atlas Okolja for their own analyses.

Figure 26. Satellite picture of Ljubljana with flood risks highlighted in red, orange and yellow, created on Atlas Okolja



What data does Atlas Okolja use?

Atlas Okoja uses different kinds of data that can all be connected with geographic information system data. Atlas Okolja shows information from:

- **cadastral data** (e.g. addresses, streets, property sizes),
- **environmental data** (e.g. water quality and air pollution by heavy metals),
- **climate data** (e.g. air temperatures, wind speed and plant phenology),
- **infrastructure data** (e.g. annual average daily traffic per road),
- **water data** (e.g. flood hazard maps and information about the groundwater),
- **nature data** (e.g. protected areas and Natura 2000 zones),
- **land and soil data** (e.g. soil pollution),
- **earthquake data** (e.g. earthquake catalogue and maps of seismologic risks).

Part of the data used for Atlas Okolja is collected and published by the Slovenian Environmental Agency, but for large parts of the data the Slovenian Environmental Agency needs to reach out to other Slovenian agencies and ministries. The cadastral data, for instance, is collected by the Slovenian Surveying and Mapping Authority and the information about flood risks is created by the Slovenian Water Agency.

The data received from other agencies is delivered in a standard format to minimise the time spent on combining all different datasets.

All data used is openly available, but the website currently does not have a download functionality. To actually download the data presented on the satellite picture, you need to visit the website of the government agency or the Slovenian open data portal.

How does Atlas Okolja create an impact?

Atlas Okolja serves citizens, policymakers, insurance companies and many other individuals and organisations. **Currently Atlas Okolja has 1 200 visitors per day**.

The Atlas does not have a dedicated feedback area, but it does receive emails regularly informing them about usability issues that can be improved on.

Future updates are already in the works: in a new version users will be able to share the map they created with others via a URL and more precise measurements of elevation data per square meter will be added to create an even more detailed picture of Slovenia.

Plume Labs: the French start-up making air quality information accessible and empowering

Plume Labs in a nutshell.

- **Service:** Plume Labs is a French start-up recently acquired by AccuWeather that uses open data to forecast air quality globally.
- **Sector:** environment, health.
- **Country of origin:** France.
- **Data sources:** open data from local authorities, EU institutions and further sources.
- **Number of employees:** 22.
- **Website:** plumelabs.com

Plume Labs is a start-up founded in 2014 by the French engineers Romain Lacombe and David Lissmyr to raise awareness about air pollution as a global health threat and help fight against it. With this aim in mind, **Plume Labs focuses on bridging the environmental information gap by making more reliable and accurate information and forecasts on air quality available globally**. The start-up, which mainly only relies on open data to develop its products, has been recently acquired by the American company AccuWeather, a former client specialised in weather forecasting service worldwide.

What services does Plume Labs offer?

The starting point for Plume Labs is that **reliable air quality information is not easily accessible**. In fact, quality monitoring has traditionally been driven by local governments and agencies, leading to air quality monitoring stations being unevenly distributed around the world and therefore to many individuals being excluded from this key information. At the same time, **local data on air quality is often hard to gather and to interpret**, which further challenges the ability of citizens to be properly conscious of their local situation and makes it difficult for governments to take action.

To respond to these availability and accuracy challenges, Plume Labs has focused on the development of air quality sensors on the one hand, and on the forecasting of air quality in major cities through Copernicus data and other open data sources on the other. More specifically, Plume Labs currently offers three main products:

1. a **marketable device** able to provide users with an estimation of their exposure to air pollution;
2. a **free-of-charge and ad-free mobile application** showing the level of air pollution based on users' choice of location;
3. an **API** that businesses interested in air quality data for their commercial products and evaluations can buy to access data in an automated manner.

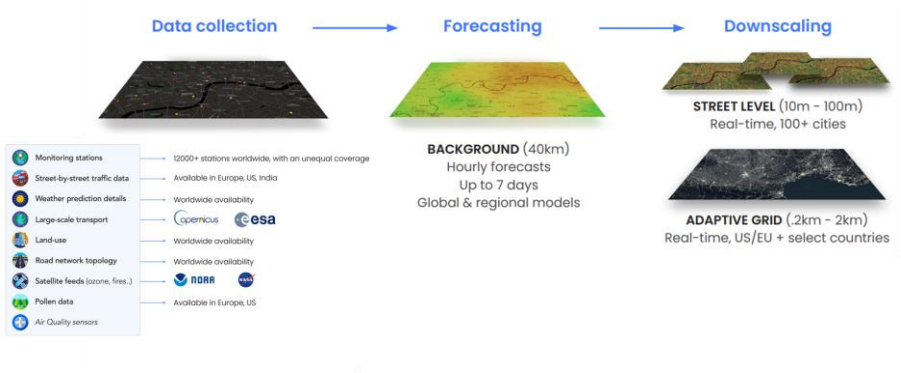
What data does Plume Labs use?

Plume Labs' products are almost completely based on open data, except for traffic data that comes from a private data provider. Among the open data, many different sources are leveraged, including local authorities for air quality monitoring stations, Copernicus API for global and EU air quality forecasts, the National Oceanic and Atmospheric Administration (NOAA) for weather forecasts, OpenStreetMap for road network and classification, and the EU open data website for urban areas and

land use. This data – which is updated hourly – refers to parameters of air quality such as wind speed and direction, precipitation, temperature, humidity, heating emissions, traffic, land use and pollen.

Based on this data, **Plume Labs trains geospatial AI and machine-learning models to produce the most hyperlocal maps and forecasts of air quality across cities around the world**. More specifically, after collecting the relevant data from the various sources, Plume Labs' team uses datasets based on historical data to create the model and inject the most recent data into the model to predict the air quality level. Subsequently, a downscaling of coarse resolution forecasts is performed up to the street level. This allows the focus on relevant air quality information to be improved and the location to be as precise as possible, which would not be possible without the open data.

Figure 27. Plume Labs' forecasting process



How does Plume Labs create an impact?

In the past 8 years, Plume Labs has been able to expand its air quality forecasting **from 60 cities to the entire world**. Currently, the start-up has **dozens of direct clients** and through AccuWeather has acquired **hundreds of additional indirect client organisations**. In terms of end users, Plume Labs also seems to have a good impact, with the mobile application product rating at **4.2 and 4.7** out of 5 on the Play store and the App store, respectively. In addition to this, given the high levels of accuracy – proven through several comparisons of performance with competitors and an ad hoc accuracy evaluation framework – Plume Labs' products are also **helping to advance the scientific research on air pollution** of major institutions such as Harvard University, University of Cambridge, Columbia University or Imperial College London.

Thanks to the recent acquisition by AccuWeather, in the coming years, Plume Labs expects to **further strengthen its global reach and extend its expertise on new forecasting models based on open data**. For instance, the team plans to develop a forecast team model to predict **burning fire evolution** in space and time. In parallel, the start-up also envisages forecasting the **wildfire smoke dispersion** to be able to trigger alerts and warnings to potentially affected people.

Baltazar: helping Croatians and tourists in Croatia find swimming spots with clean water

Baltazar in a nutshell.

- **Service:** Baltazar measures water quality on beaches in Croatia. The data is further enriched with information about air temperature, wind speed and beach facilities.
- **Sector:** environmental.
- **Country of origin:** Croatia.
- **Data sources:** open geodata.
- **Number of employees:** n/a.
- **Website:** vrtlac.izor.hr

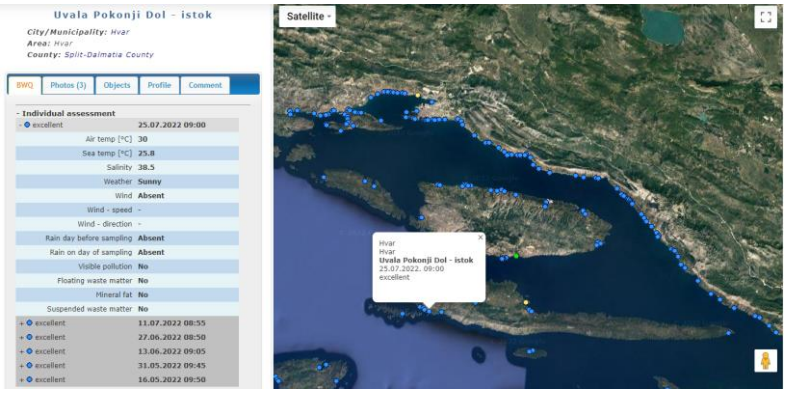
The application Baltazar was developed to **provide insight into the water quality on beaches** (e.g. if water in certain areas is suitable for swimming) **and to increase awareness of the connection between environmental sustainability and water quality**. Moreover, Baltazar provides important information to beachgoers about the facilities on beaches in Croatia. Lastly, the data collected by Baltazar is published as open data to enable reuse by NGOs and private organisations.

What services does Baltazar offer?

Baltazar was launched by the Croatian government to better inform citizens and tourists about the water quality around the Croatian beaches. The goal is twofold: to create a water quality database to comply with the Croatian directive on sea bathing water quality that is also suitable for reuse; and to **provide an easy-to-use web interface where users can find regularly updated information about sea water quality**. Recently, Baltazar also created a mobile app so users can easily find the water quality of beaches on their smartphones.

Baltazar contains data for over 1 000 beaches. The water quality is colour coded as either excellent, good, sufficient or poor. Baltazar also provides information about the air temperature, sea temperature and the profile of the beach, for instance. Moreover, it provides data on parking spots, hotels, nearby restaurants and rubbish bins. In short, everything that is needed for a nice day on the beach.

Figure 28. Water quality and other information about one of the Croatian beaches



What data does Baltazar use?

Baltazar uses open geo data to show the beaches on a map. The data on water quality is collected by microbiological researchers that regularly visit each of the more than 1 000 locations and is sampled every 2 weeks. Furthermore, open data from meteorological stations is used to show the air temperature and wind speed on specific beaches.

Other available data on Baltazar, such as the location of rubbish bins and the availability of parking spot, is collected by the Baltazar team by zooming in on satellite photos to identify objects (i.e. rubbish bins), and by users who can share photos and comments about a beach on the website.

Baltazar is currently the central point for data provision of water quality data in Croatia. **The data they collected and enriched with open data is used by several NGOs to evaluate the water quality** in Croatia and by private organisations to create services that inform users about beach facilities.

How does Baltazar create an impact?

Baltazar helps citizens and tourists choose safe places to swim, but the precise positive impact that Baltazar has on public health has not been estimated. The website was used by **over 260 000 unique visitors between May and August 2022**, of which 244 000 used the website in Croatian and 16 000 used the English version.

Baltazar uses the feedback they receive from users to keep improving their services. For instance, users often suggest new locations to be evaluated. Baltazar assesses these requests and sometimes adds the new beach to the data. Users can also comment which features are not working properly on the website, which can then be addressed by the Baltazar team.

Environ-Mate: the German platform empowering kids with knowledge about climate change

Environ-Mate in a nutshell.

- **Service:** Environ-Mate is an interactive platform to empower children with knowledge about climate change based on scientific data.
- **Sector:** environment, climate.
- **Country of origin:** Germany.
- **Data sources:** open data from EU and international bodies (e.g. Eurostat, the European Environment Agency (EEA), NOAA).
- **Number of employees:** 4.
- **Website:** environ-mate.feld-m.de

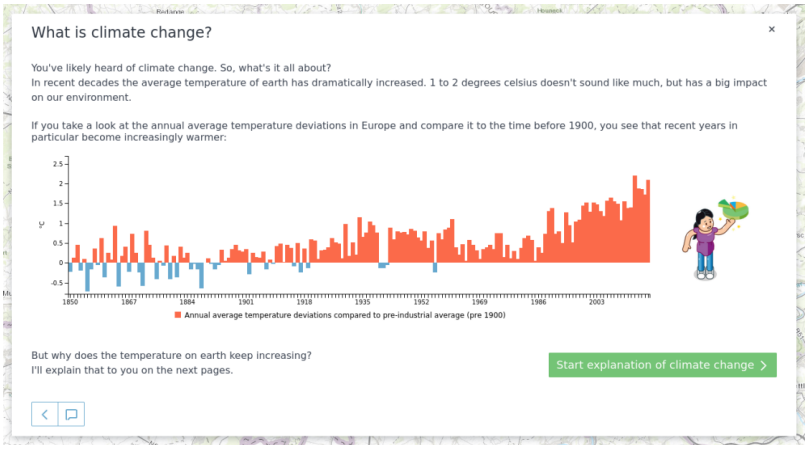
Environ-Mate was developed in 2019 by Dr Matthias Böck, Bernhard Janetzki, Alexander Merdian-Tarko and Paul Schlumbom. To respond to the 2019 EU Datathon challenge 'Tackling climate change', the German team developed an interactive web-based application that **explains climate change to children between 10 and 14 years old and encourages them to become engaged in fighting against it**. The application is completely based on open data from the EEA, NOAA, Eurostat, Copernicus and the Centre for Research on the Epidemiology of Disasters emergency events database (CRED EM-DAT), and its source code is available on Github.

What services does Environ-Mate offer?

Aware of the devastating consequences of climate change, leveraging the huge amount of data about it and in light of the growing activism among young people, the Environ-Mate team decided to create a **web-based application able to educate children about climate change and the means to act**. Today, Environ-Mate is an interactive and intuitive platform that uses scientific data to teach children and young adults about the causes and effects of climate change.

More specifically, the user is accompanied on a journey about **various statistics in reference to their own home country in Europe**. The app provides information on: what climate change is exactly, what types of greenhouse gas (GHG) emissions exist, where these gasses come from, how European countries are doing in comparison to each other, what the main consequences of climate change are and what individuals can do to prevent it. To sum up, Environ-Mate is guided by a double principle of learning and acting.

Figure 29. Environ-Mate: 'What is climate change?'



What data does Environ-Mate use?

The platform offers recent and reliable information on climate change in a specific country and the whole of Europe. This **data is completely open** and comes from the EEA, NOAA, Eurostat, Copernicus and CRED EM-DAT.

Besides the cleaning and harmonisation of the data, the team also used a digital elevation model and GHG projections to provide insights on the annual average temperature increase, GHG emissions per country and sector, the comparison of GHG emissions across Europe per capita, total and per sector, the GHG emissions per country over time, the sea level rise forecast map, the extreme weather occurrences and the projections on temperature increases.

How does Environ-Mate create an impact?

The Environ-Mate team has not implemented any tracking in its web application to be compliant with legal and privacy regulations, especially since the target group consists of children. Yet, the platform was evaluated through qualitative surveys and received positive feedback from several individuals, including during the 2019 EU Datathon. In general, the idea behind Environ-Mate appears to provide some value to the target group although there is certainly room for improvement considering the recent fierce competition on the internet.

For the future, the team has some ideas for further development and the next steps. For example, they would like to **test Environ-Mate in schools and learn from this feedback, do a climate expert audit, optimise accessibility features for people with disabilities and extend stories** (e.g. by providing more details on biodiversity, region-specific topics and having a greater global coverage). They are also considering **making the platform available in more languages** (French, Italian, Polish and Spanish). The implementation of these ideas will, however, depend on the time and resources of the team members next to their professional and private responsibilities.

4. Learnings from the four dimensions

Economic impact

Reuse cases create economic impact in different ways. The organisations **C4P and YouthPoP** use data from the Commission's TED portal. Currently, some organisations struggle to find tenders that suit their needs and qualities on the portal. C4P and YouthPoP analyse the open tendering data and present the most promising opportunities to their clients. In doing so, they established a business model for themselves, while improving the match between organisations looking for opportunities and the EU institutions that put the tender on the market.

Wonder Wanderlust Women and Naar Jobs in West-Vlaanderen both use open data to help people find a job and place that suits them, although their target groups are different. Wonder Wanderlust Women aims to support young women in their personal and professional growth, whereas Naar Jobs in West-Vlaanderen helps people that are currently unemployed in finding a job that they can reach with the transport options available to them. Together, the initiatives show how open data can bridge gaps in the labour market that are currently left unaddressed.

Lastly, **LocalFocus** curates datasets for journalists. LocalFocus collects the open datasets that are most newsworthy and presents these datasets to journalists. This increases the impact of these open datasets substantially because journalists sometimes have trouble finding their way in all available open datasets. In the data life cycle, they play the part of the open data intermediary. LocalFocus also offers its users multiple analyses per month and provides them with a visualisation tool to further increase the journalistic impact of their stories.

Governmental impact

Reuse cases that have a governmental impact often improve political transparency in countries. For instance, the Norwegian **statsregnskapet.no** is part of a government agency and collects open data from multiple government sources to create visualisations of government spending, informing taxpayers on where their money goes. **Openpolis** provides several similar services. For example, they provide open insight into the balance sheet of Italian municipalities and updates of the implementation of the Italian recovery and resilience plan. The Georgian initiative **IDFI** also makes governmental data available to the public. Additionally, they also pressure their government into making more data open. The idea is that only a transparent government can be held accountable.

Another example of governmental impact is **Where is my polling station?** as it helps people find a voting booth close to them that suits their needs regarding accessibility. Open data is thus used to broaden political participation and in turn strengthen democracy. A similar example is **Next Generation Democracy**, which keeps a democracy technology database, develops an online ranking of the performance of MEPs and continuously identifies opportunities to leverage technology for more democratic societies.

By using network analysis to visualise the evolution and interconnectedness of case-law, **The Smartfiles Network** not only revolutionises the world of PDFs, but also makes legal text and key court decisions easier to grasp by citizens and to act on by policymakers. In a similar way, the **3D city model** of the city of Aarhus makes use of data-driven visualisation technology to facilitate the decision-making process of the municipality when it comes to climate adaptation, green conversion, urban planning, land management and much more.

Social impact

Open Food Facts shows how opening up data can lead to real impact. Their open data (of over 2 million products) was partly created by foodies that would upload product labels with the ingredients. Open Food Facts popularised the nutri-score (i.e. an easy-to-understand number that tells you the nutritional value of a product) in France to the point that large food producers cooperate with them to make their products healthier and improve public health.

Other reuse cases that contribute to public health are: **Medicatio**, which provides a database of approved drugs that patients can use to better understand the care they're receiving; the application **EVapp** that shares the location of people suffering from a cardiac arrest with the closest first aiders, thus decreasing the chance of death or brain damage; **Hale & Hearty** that uses open data to encourage Irish citizens to start living a healthier lifestyle; **OpenActive**, which created an easy way to find nearby sporting locations; **VisImE-360** that helps to raise awareness and improve decision-making in relation to visual impairment; and **UniversiDATALab** that offers an interactive repository of analytical applications of research in higher education in Spain.

Tangible data uses open data to let people experience the data. They want to make large data sculptures that can be placed in front of government buildings instead of complex dashboards that are only understood by few. Tangible data's impact is to challenge fake news and to provide accurate information to people with little digital skills.

In a similar way, **EU Twinnings** uses open data in an exploratory way, allowing a wider audience to access Eurostat statistics and showing the degree of similarity between EU regions using interactive graphs.

Lastly, the reuse case **Integreat** provides information about German municipalities in several languages to help newly arrived migrants and refugees to integrate into the new society.

Environmental impact

Multiple reuse cases used open data to monitor the air quality and inform (vulnerable) citizens accordingly. Nevertheless, what they measure differs: **Air Quality Cyprus** measures pollution from industries or nearby deserts; **Planttes** monitors which plants are in bloom; **Plume Labs** uses satellite data from Copernicus to predict air quality in several cities; **Baltazar** uses data on air temperature and wind speed to measure water quality on beaches.

Open data plays an important role in preserving nature and protecting the planet. **Digital Forest Dryads** tracks illegal deforestation with open aerial data and **Vides SOS** can be used to alert the Latvian State Environmental Service of environmental hazards such as pollution, waste and litter in nature. **Atlas Okolja** informs Slovenian citizens about noise pollution, air pollution, earthquakes and Natura 2000 areas in their country.

Finally, the environmental impact of **Environ-Mate** passes through the 10- to 14-year-old children that the platform tries to educate on matters pertaining to climate change and the means to act.

5. Conclusion

The use case observatory monitors 30 reuse cases over the course of 3 years to assess how impact is created with open data, to share the challenges and achievements of open data reuse cases and to add to the debate on an open data impact assessment methodology.

An extensive procedure was followed to select the most relevant reuse cases for this research project. Drawing from the ODM assessment, the EU Datathon and the data.europa.eu use case repository, 600 interesting reuse cases were collected. A relevant sample was chosen to balance out the number of reuse cases between different countries and across different impact domains (economic, governmental, social and environmental). The last criterium was whether the reuse case benefitted disadvantaged groups in society or contributed to the Commission's policy priorities for 2019–2024.

The interviews carried out with the developers of the reuse cases clearly show the myriad of ways in which open data reuse cases can have an impact. For example, with open data: new business models are established; people are helped to find jobs close to them; citizens can better understand how their MEPs and their government operate through improved transparency; people suffering a cardiac arrest see increased survival rates; and children are empowered with knowledge about climate change.

At the same time, estimating the impact of these reuse cases is a complex challenge. For many reuse cases it remains unclear how to precisely measure their services' impact. An organisation like Open Food Facts contributes to public health by providing citizens with simple-to-understand information on the nutritional values of products. However, how much they contribute to public health – which is often measured in QALYs – would require extensive research. Therefore, most of the reuse cases measure their impact with web statistics – the number of unique visitors serves as a rough proxy for the impact that is created by the reuse case.

Another challenge in estimating the impact of open data reuse cases is that not all information can be shared easily. Sharing precise information about the revenue or the number of clients could be a risk for SMEs that use open data in their business model, since competitors could use that information to their advantage. This is different from public sector institutions and non-profit organisations, who are freer in sharing information that is available to them.

Each interviewee was also asked what plans they have for the near future. In other words: how will they continue to create an impact? The ideas in this regard differ from only small changes to impressive ambitions. However, a common challenge for both small and big plans seems to be finding appropriate investments. While finding the resources to create the reuse case or a first prototype is feasible, obtaining the funding to scale the reuse case is not so straightforward. Therefore, some of the use cases are struggling to keep moving forward.

Overall, this first report highlights the power of open data reuse. Without open data, most of the organisations and applications mentioned in this report would have never existed. At the same time, however, the results indicate the need to further unlock the potential of open data, which would allow its reuse to have a clear impact on our economy, government, society and environment. While this requires further supporting the community of reusers in identifying financial opportunities for growth, it also entails developing a better understanding of how open data impact is created and measured. The observatory aims to achieve this understanding, and with this first report it has laid the foundations for continuing this journey. The idea is to use the insights of the first report as a benchmark for a second and third analysis that will be published in 2024 and 2025. With respect to this first report, which introduces the 30 reuse cases monitored, the second and the third report will focus more on

evaluating the progress made. Moreover, more emphasis will be placed on ascertaining achievements and challenges in a period of 3 years and extrapolating concrete insights to improve methodologies of open data impact assessments.

Annex I – List of interviewees

No	Reuse case name	Interviewee	Contact
1	the Smartfiles Network	Veronika Haberler	vh@lereto.at
2	3D city model	Marianne Knudsen	maknu@sdfi.dk
3	Waar is mijn stemlokaal	Jesse Renema	jesse@openstate.eu
4	OpenPolis	Vittorio Alvino	fondazione@openpolis.it
5	Next Generation Democracy	Michael Birkebæk Jensen	nextgenerationdemocracy@gmail.com
		Kristian Thorsted Madsen	kristian.t.madsen@gmail.com
6	Statsregnkapet	John André Jakobsen	john.andre.jakobsen@dfo.no
7	IDFI	Teona Turashvili	t.turashvili@idfi.ge
8	Medicatio	Willy Duville	willy.duville@medicat.io
9	UniversiDATALab	Juan Jesús Alcolea Picazo	jjalcolea@dimetrical.es
		José Arbues Bedia	jarbues@ucm.es
10	ViSimE-360	Boris Bikbov	boris.bikbov@gmail.com
11	Open Active	Tim Corby	hello@openactive.io
		Howard Askew	
12	Tangible Data	Antonio Moneo	antoniomoneo@gmail.com
13	Hale & Hearty	Helena Campbell	Helena.Campbell@per.gov.ie
14	EU Twinnings	Giuseppe Sollazzo	puntofisso@gmail.com
15	OpenFoodFacts	Stéphane Gigandet	stephane@openfoodfacts.org
16	Integreat	Clara Barcklo	bracklo@integreat-app.de
		Fritjof Knier	knier@integreat-app.de
17	EV-app	Robin Arys	robin.arys@evapp.org
18	Digital Forest Dryads	Razvan Pistolea	razvan+europa@digital-dryads.eu
19	Air Quality Cyprus	Chrysanthos Savvides	csavvides@dli.mlsi.gov.cy
20	Vides SOS	Liene Alde	liene.alde@vvd.gov.lv
21	Planttes	Jordina Belmonte Soler	jordina.belmonte@uab.cat
22	Atlas Okolja	Primož Kogovšek	primoz.kogovsek@gov.si
23	Plume Labs	Boris Quennehen	boris.quennehen@accuweather.com
24	Baltazar	Damir Ivankovic	ivankovic@izor.hr
25	Environ-Mate	Alexander Merdian-Tarko	alexander.merdian-tarko@posteo.de
26	C4P	Andreas Thanopoulos	andreas@c4p.io
27	WWW by ITER IDEA	Sara Baroni	sara.baroni@iter-idea.com
		Guido Mazza	guido.mazza@iter-idea.com
28	YouthPop	Sofia Lousa	sofia@youthpop.eu
		Michail Maragakis	melidoni.michail@gmail.com
29	LocalFocus	Jelle Kamsma	jelle.kamsma@anp.nl
30	Naar Jobs in West-Vlaanderen	Han Tambuyzer	han@nazka.be

Annex II – Indicative interview questions

1. Can you briefly describe the **idea** behind your use case?
    a. What is its purpose?
    b. What is its target audience?
    c. What size is the team?

2. Can you briefly describe the **data** that was used for creating *x*?
    a. Is your initiative only based on publicly accessible data (open data)?
    b. If not, which other data sources were consulted?
    c. Where did you find the (open) data?
    d. What is its licence?
    e. How do you process open data? Any cleansing, structuring, manipulations or modelling techniques?
    f. Would it have been possible to develop your initiative without open data?
    g. If not, why?

3. Can you briefly describe the **impact** that *x* is having on its audience, with respect to its purpose?
    a. Do you monitor the performance of your website? For example, do you know the number of daily/monthly active users?
    b. Do you regularly collect feedback from users? Is this feedback proving a general satisfaction of users with the website?
    c. How has *x* been developing since your win/participation in the EU Datathon of 2020?
    d. How do you plan to further develop from now to the next 2 years? Is there any particular ambition or business projection that you would like to share with us?

4. Is there any **further information** that you would like to share with us?

ISBN 978-92-78-43364-2