ESSnet Big Data II

Grant Agreement Number: 847375

https://webgate.ec.europa.eu/fpfis/mwikis/essnetbigdata
https://ec.europa.eu/eurostat/cros/content/essnetbigdata_en

Workpackage L

Preparing Smart Statistics

Deliverable L2: Description of the findings regarding Task 2, Smart Cities

Final version, 1st November 2019

Prepared by:

Case Study 1:
Mimirium Ltd. (sub-contractor of the BNSI, BG)
Deyan Slavov and Galya Stateva (BNSI, BG)

Case Study 2:
Natalie Rosenski, Heike Kürten (Destatis, Germany)
Dr. Holger Leerhoff, Nicole Jurisch (SSO BB, Germany)
Massimo De Cubellis, Giovanna Astori (Istat, Italy)

Case Study 3:
Marie-Pierre de Bellefon (INSEE, FR)
Massimo De Cubellis (Istat, Italy)

Workpackage Leader:

Natalie Rosenski (Destatis, Germany)
Mail address: natalie.rosenski@destatis.de
Telephone: +49 611 754284

Executive Summary

Considering the urban growth during the last 20 years, European cities have become a highly complex system with a rising number of urban challenges such as traffic congestion and air pollution, energy consumption and fuel poverty, economic growth and job creation as well as creating a citizen oriented city council by responding to citizen needs. In course of the megatrend digitalization, the availability of IoT (Internet of things)[1] tools such as connected sensors offer the possibility to measure on a continuous basis the extent of emissions, energy consumption and mobility. By using sensors and a digital infrastructure with urban data platforms, cities turn into smart cities in order to improve the quality of life of citizens and businesses, by increasing sustainability and by decreasing emissions e.g.

In task 2 of WPL, the National Statistical Institutes (NSIs) of Bulgaria (BNSI), Germany (Destatis and the State Statistical Office of Berlin-Brandenburg, SSO BB), France (INSEE), Italy (ISTAT) and the United Kingdom (ONS) investigate the potential of different IoT technologies that are used in the context of smart cities (definition see below) in order to produce trusted smart statistics. The task explores the currently used smart technologies in the context of smart cities in two different perspectives: how could NSIs benefit from the collected data of the used IoT tools such as smart sensors and is there potential in the context of the smart cities for producing smart statistics? In other words, task 2 examines how NSIs could support cities to transform themselves into sustainable and participative smart cities in three specific case studies.

Case study 1 is an example of a collaboration between the Bulgarian NSI and Mimirium Ltd., a company which sets up technological solutions for monitoring individual's mobility. The goal of this partnership is to explore the existing IoT infrastructure and technologies in the pilot city of Varna and to define possible ways of using this infrastructure for smart statistics purposes. The implementation of a pilot-project (Varna limitless) demonstrates the practical feasibility of these ideas. Varna Limitless is an ecosystem of different tools: a mobile application, anonymization nodes, and a web-interface, which monitors the movement of diverse groups of users with different disabilities and compares them with the movement of a control group of users without disabilities. Thanks to the data collected, the municipality of Varna will be able to detect areas which are particularly difficult to reach for disabled people, in order to focus their efforts on improving accessibility to these parts of the city. The project will also help to build a more active and involved local community through their participation in the project. All collected data are to be accessed by the NSI, which makes them a trusted smart statistics source.

One of the challenges of this project is to guarantee the confidentiality of individual mobility data. For this purpose, the partner Mimirium Ltd. used the existing IoT infrastructure and a very efficient blockchain technology. This successful experimentation can now be expanded with more volunteers and tested in other cities. More generally, this pilot project proved that some technological solutions such as blockchains can minimize legal risks regarding data protection and convince individuals to provide an NSI with detailed data on their mobility patterns.

[1] The Internet of Things is known as a system of interrelated computing devices, such as sensors, but also machines, objects etc. that have the ability to transfer data over a network without requiring human-to-human or human-to-computer interaction (see https://en.wikipedia.org/wiki/Internet_of_things, accessed on 31 October 2019).

In case study 2, the three partners Destatis, ISTAT and SSO BB examined the exemplary cities Berlin and Darmstadt and 12 projects that are funded by the Horizon 2020 (H2020) programme of the European Commission supporting smart cities and communities. The case study aims at giving a first overview of the different solutions, which selected smart cities in Europe already have or are planning to implement in order to improve the services and infrastructure of the cities for the benefit of its inhabitants and businesses.

A variety of interesting data sources is given by listing the most promising solutions for official statistics used in each city or rather H2020 project. The data is sorted by the identified domains energy, mobility and information and communication technologies (ICT). There are several promising types of data available in the smart city context, including (but not limited to) sensor data on energy production and consumption on various levels and different kinds of fine grained sensor data on traffic and (smart) mobility. Urban data platforms could serve as means for collecting and preprocessing these sensor data and provide an interface-/access point-functionality for official statistics. These sensor data can complement the existing energy or mobility statistics or could be a first step for creating new distinct statistics. Nonetheless, the availability and the access to these promising data sources are currently rather poor. This might be due to the fact that some of the projects are still ongoing, but also because some cities might not want to publish their data on public websites. Concerning the data needs it turns out that smart cities generally require data on a small-area level, ideally on a district or even smaller level. Rather often these data –if even available – can hardly be provided by official statistics due to confidentiality reasons. Despite these challenges, three specific use cases are suggested in this case study, which are promising for a follow-up study.

The first suggests a collaboration with the city of Darmstadt to further explore the data available there, especially the real-time traffic data that is already available online for free. The second use case refers to the topic 'data platform' that potentially includes all data that is produced in smart cities. The third use case indicates to further explore 'smart lamp posts', which are collecting mobility and environmental data e.g.

In case study 3, INSEE and ISTAT explored how combining data - collected by air pollution sensors, NSI's socio-economic data and city data about mobility traffic - could guide local environmental policies.

INSEE works together with IMREDD (Mediterranean Institute for Risks, Environment and Sustainable Development) on a first study about the links between socio-economic vulnerability and exposure to air pollution in the city of Nice. This study has proven that in the city of Nice, people of higher socio-economic status are significantly less exposed to air pollution than people of medium and low socio-economic status. Based on these results IMREDD works on an analysis to understand the reasons of high air pollution in certain areas. IMREDD therefore investigates information on car traffic in the most underprivileged areas. By using IMREDD's mobility modeling software and INSEE's data about the working place of Nice's inhabitants, commuting simulations can be evaluated. Eventually, since IMREDD works jointly with the Nice Metropolis, which is responsible for the installation of pollution sensors, the Nice Metropolis will use this study to identify areas, where a pollution sensor should be placed in order to make a more precise analysis. Therefore, this is a very promising collaboration between different institutes and the city council, which on the one hand is in charge of data collection and on the other hand the public policy maker using the data.

In the medium term, this study aims to convince cities how fruitful the sharing of data with NSIs can be to enrich their local policy decisions. Therefore, the analysis should be reproduced in other cities and the results could be compared to other cities. Another axis of improvement for the medium term is to study the socio-economic characteristics of people exposed to pollution during day time. Meanwhile, ISTAT will exploit the experience gained by INSEE colleagues and continue in defining the partnership with the Regional Agency for Environmental Protection.

These three case studies show the most current usage possibilities of new digital data for smart cities and give approaches for further work: All projects within these case studies are planned to be replicable for other cities with similar profiles and issues. Although, before diving into the three case studies, a common definition of smart cities will be given in the following.

Definition of Smart Cities

Since the term is still quite new in data science and politics, no explicit definition of a smart city exists yet. According to Wikipedia[2], "a smart city is an urban area that uses different types of electronic IoT sensors to collect data and then use insights gained from that data to manage assets, resources and services efficiently. This includes data collected from citizens, devices, and assets that is processed and analyzed to monitor and manage traffic and transportation systems, power plants, utilities, water supply networks, waste management, crime detection, information systems, schools, libraries, hospitals, and other community services."

In line with this is the definition of the European Commission[3] that describes "a smart city [as] a place where traditional networks and services are made more efficient with the use of digital and telecommunication technologies for the benefit of its inhabitants and business." Also here, it goes beyond the use of ICT, but rather "means smarter urban transport networks, upgraded water supply and waste disposal facilities and more efficient ways to light and heat buildings. It also means a more interactive and responsive city administration, safer public spaces and meeting the needs of an ageing population." According to the H2020 programme technological and service solutions to reach these aims require "integrated approaches, both in terms of research and development of advanced technological solutions, as well as deployment."

[2] https://en.wikipedia.org/wiki/Smart_city
[3] https://ec.europa.eu/info/eu-regional-and-urban-development/topics/cities-and-urban-development/city-initiatives/smart-cities_en

Table of Contents

Case Study 1

1 Introduction

In task 2, case study 1 of WPL, the National Statistics Institute (NSI) of Bulgaria (BNSI) has explored the possibilities of building sustainable projects for smart statistics in Varna, Bulgaria, in partnership with Mimirium Ltd. We explored the existing Internet of Things (IoT) infrastructure and technologies in Varna and the possible ways of their use for smart statistics purposes and execution of a small pilot project (Varna Limitless). We built a pilot project based on the existing resources and gathered some statistical data. We proved that it is possible to collect data from IoT and our next goal is to enlarge the amount and type of data we collected and processed.

Our first task was to explore the existing IoT infrastructure and technologies, possible ways of their use for smart statistics purposes and to execute a small pilot project (Varna Limitless) in order to verify these possibilities in the best practical way. Varna Limitless is an ecosystem of different tools: mobile application, anonymization nodes and web-interface, which monitors the movement of different groups of users with different disabilities and compares them with the movement of a control group of users without disabilities. Then we can identify the problem areas as well as the movement patterns for each group and compare the differences. Once the models are created, personalized polls will be sent to individual users (respondents taking part in the pilot project) to gather detailed information about the issues people are experiencing in the detected problem areas.

The goal is to do some good for the community and to develop better local facilities and services. By utilizing this project, we will help to build a more active and involved local community through bringing them in the project as a control focus group that will be used as a measurement of the other groups' problems in moving through the city. In this regard, BNSI was also able to gather and provide useful statistical data that can be used as a trusted smart statistic[4] source. And last but not least, we build a knowledge network within the business community that will help companies to better understand some customers and their specific needs.

During our research we explored some of the most cutting-edge technologies like blockchain, Homomorphic encryption[5], Zero-Knowledge Proofs[6], Secure multipart computation[7] that can provide the needed anonymization and aggregation to achieve a compliance with the General Data Protection Regulation (GDPR).

For the execution of the pilot-project, we selected two different groups of 10 individuals each. A group of 10 disabled people (all of them with partial loss of their movement capabilities with 25% to 50% degree of disability) and a control group of 10 people without disabilities. All of them males, between 35 and 40 years old, actively working, located in one

[4] Trusted Smart Statistics can be seen as a service provided by smart systems, embedding auditable and transparent data life-cycles, ensuring the validity and accuracy of the outputs, respecting data subjects' privacy and protecting confidentiality. (https://ec.europa.eu/eurostat/cros/content/trusted-smart-statistics-nutshell_en)

[5] A form of encryption that allows specific types of computations to be executed on cipher texts and obtain an encrypted result that is the cipher text of the result of operations performed on the plain text.

[6] Allow data to be verified without revealing that data

[7] A way that data can be used to analyze complex problems

particular district of Varna city (Levski) in order to keep the data consolidated. These restrictions were applied, because for the purpose of this pilot project we need a compact group of individuals situated in a relatively small area of Varna city. According to the data provided by 'The Union of Disabled People' in Varna, this group was the most suitable for this particular project.

At the beginning of the project we set to gather information for 5 different sets of data:

1. Common spots visited by citizens in Varna
2. Average number of visits to certain spots and places
3. Average speed of movement through the city of people without disabilities
4. Average speed of movement through the city of people with certain disabilities
5. Difference of the average speed between both groups

Mimirium with their pilot project Varna Limitless has proven that such a smart system can be achieved and built even with the existing infrastructure and technologies. All collected data can be used as a base for official statistics and can give an interesting insight and even information that previously was not possible to collect and verify. The possibilities that distributed networks and solutions can give users are fascinating and we can achieve more trust and reliability. Distributed ledgers are inherently harder to attack, because instead of a single database, there are multiple shared copies of the same database, so that a cyber-attack would have to attack all the copies simultaneously to be successful. The technology is also resistant to unauthorised change or malicious tampering, because the participants of the network would immediately spot a change to any part of the ledger. Additionally, the methods by which information is secured and updated allow participants to share data and to be confident that all copies of the ledger match each other at any time. Based on such solutions we can minimize potential legal risks such as GDPR and E-privacy, so that we can minimize data breaches and data leakages. We can provide respondents 100% anonymity and security of their data, while providing BNSI a powerful and flexible instrument for data mining of different types of data. Apart from the data, based on the GPS information, the Mimirium app and infrastructure can also be used for:

- Gathering aggregated information from distributed network of IoT sensors
- Gathering aggregated and anonymized behavioral users' data from mobile devices
- Gathering aggregated and anonymized behavioral users' data based on their internet activities, including interests, purchases, searches etc.
- Gathering aggregated and anonymized behavioral users' data based on their social network activities, cross-relating data from different sources

The possibilities are endless. In a more and more digitalized world, solutions like Mimirium Network and Varna Limitless are just the beginning of a whole new era of user-driven and user-centric apps.
Due to data privacy and reasons of trust into official statistics one need to ensure a careful handling of data with each new information or technical device. People have to be sure that they can control their digital identities and all data they generate in the global network. They have to be sure that they receive a fair reward for their time and attention too.

We live in a time and environment where we strive for the epithet ‘smart’. We are surrounded by smartphones and televisions, smart cars, smart gadgets etc. and many cities across the globe are moving forward to that prestigious and increasingly important title ‘smart city’. However, what does it actually refer to? ‘Smart city’ is an urban area that uses different types of electronic data collection sensors to supply information, which is used to manage assets and resources efficiently. This includes data collected from citizens, devices, and assets that is processed and analysed to monitor and manage traffic and transportation systems, power plants, water supply networks, waste management, law enforcement, information systems, schools, libraries, hospitals and other community services.

Cities are seeing unprecedented growth, bringing major challenges as they seek to remain sustainable, healthy and safe places for people to live and work. The IoT offers new opportunities for cities to use data to manage traffic, cut pollution, make better use of infrastructure and keep citizens safe.

2 Key drivers for IoT applications in Smart Cities

There are several reasons for municipalities to move to the wireless communications methods offered by IoT technologies.

- Cost is a key driver in the decision to convert from wired to wireless solutions, as it is massively expensive to install and maintain landlines. Additionally, the costs of cellular data plans are dropping and the robustness and throughput of wireless communications are therefore enabling new use cases that would previously have been cost-prohibitive.

- Efficiency is another important impetus. In most wired solutions, service personnel must physically go to the installation site to audit and service the communications infrastructure. These ‘truck rolls’ are expensive and inefficient, since they typically occur on a schedule, whether or not a problem exists. By contrast, wireless communication enable remote monitoring and management of IoT deployments. This enables administrators to perform firmware updates and security patches across the entire deployment, and get automated notifications in the event of any issues.
- Resource reduction is often a driver as well, particularly in use cases such as smart street lighting and monitoring assets. These IoT applications make it possible to use sensors to gather data and wireless modules to control resource use, which can result in a dramatic reduction in energy use.

3 IoT main areas in the Smart Cities of the future

Smart architecture and energy management - everything from the shape and structure of buildings to their lighting is currently being geared toward being greener and having minimal impact on the environment, especially since energy efficiency has a major effect on building value. The IoT provides exciting potential to harness energy usage in a more efficient manner. Street lighting and meter systems for resources such as water, gas and electricity are also being updated. A recent Amsterdam program[8] installed smart electricity meters in homes to allow for automatic management of generated solar power and the sale of excess electricity to the central grid.

Security and privacy - the most obvious area, in which the IoT has been implemented to improve security in urban areas, is video surveillance, which allows for the police to monitor live feeds from across an entire city, relying on Artificial Intelligence (AI) systems to detect and report incidences of crime or the sighting of wanted persons or vehicles. As city administrators and tech developers work to strike a balance between privacy and efficacy, some programmes have already been implemented including a gunshot detection technology in some parts of New York City, which automatically alerts nearby police officers whenever a gunshot is captured.

Traffic management - as cities have expanded and affluence has spread, the number of cars and drivers has clearly increased, which comes with the inevitable problem of congestion and higher incidence of accidents. Several innovations have been made to solve the problem, from redesigned traffic lights to improved parking systems, but it is still a major issue in most cities. The first way the IoT is addressing traffic problems is by replacing the current system of fragmented and unreliable self-reporting by drivers with a comprehensive system that gives accurate, real-time information about traffic dispersion, making it easier for drivers to be guided to other roads or for autonomous vehicles to simply reroute on their own. Parking automation is another developing area. In Shanghai, a pilot program[9] is being run by the city to allow drivers to search for, book, and navigate their way to a parking space as well as pay directly with their smartphones. Automation like this reduces the loss of productivity while people circle blocks in search of space and provides extra revenue.

Ultimately, there is massive potential to innovate and make profit from the shift toward the IoT across various sectors, including city planning. The difference between entrepreneurs who will succeed in the IoT industry and those who will not is whether their innovations maximize the features of the IoT to improve efficiency and reduce environmental impact while keeping privacy rights strong.

Short description of Varna Limitless Project

Everyday life of disabled people in Varna is full of challenges in order to function as normal members of the community. One of the main pains for disabled people is the public infrastructure. It is difficult to access specific areas and buildings, because of the lack of well thought out infrastructure.

That’s why, at Mimirium, it was decided to design a solution called Varna Limitless for improving the infrastructure of the 3rd largest city in Bulgaria — Varna. Varna is the Sea Capital of

[8] The Netherlands smart meter rollout continues to draw interest as a good test case for how to digitise electricity and gas measurement at scale - https://www.engerati.com/article/dutch-smart-meter-experience-lessons-mass-rollout

[9] https://www.scmp.com/news/china/society/article/3011170/apps-helping-shanghai-tackle-its-chronic-car-park-shortage

Bulgaria and the population grows to 750.000 people in the active season. Through the last years a lot of improvements were made in the city's infrastructure, but even in a big city like Varna, we still see every day how disabled people face a lot of challenges while moving through the city.

Our team began thinking about how we can measure these difficulties. We needed quantification in order to identify and analyze the problems. In order to solve the accessibility problem, first we needed to gather data that shows, which are the hot spots and bottlenecks in the city.

As a result, we decided to create an application, which monitors the movement of different groups of users with different disabilities and compare them with the movement of a control group of users without disabilities.

Then we can identify the problem areas as well as the movement patterns for each group and compare the differences. Once the models are created, personalized polls will be sent to individual users to gather detailed information about the issues people are experiencing at the detected problem areas.

Information is collected by GPS and additional installed beacons at the crucial points in the city.

But who would sacrifice their own privacy and allow to be tracked for the sake of a greater good?

Our experienced team utilizes modern technologies to achieve complete anonymization of all participants. We will achieve this, by storing the gathered data locally in an encrypted model on users' devices. The transmission to the system will happen only after anonymization and aggregation of the information has taken place — so only processed information will be provided by the users of each group. Everyone will remain anonymous and the privacy is fully protected.

Our system uses innovative cryptographic practices and blockchain technology to fully guarantee the security of personal data. We use homomorphic encryption, zero-knowledge proofs, secure multipart computations and Etherium blockchain in order to provide high level of security for our users. All of these technologies are desribed below.

It will also allow participants involved in the project to receive payments for the data they have generated.

So, would you participate if you knew that:

- you are supporting a good cause — better infrastructure in the city for all people;
- you are totally anonymous by helping;
- you are rewarded with money for helping.

It's a win-win situation.

What is the project goal?

The goal is to make something good for the community and to develop better local facilities and services. By utilizing this project we will help to build a more active and involved local community through bringing them in the project as a control focus group that will be used as a measurement of the other groups' problems in moving through the city. Together with the National Statistics Institute (NSI) of Bulgaria our team can gather and provide useful statistical data that can be used as a trusted smart statistic source. And last, but not least, we build a knowledge network within the business community that will help all companies to better understand all their customers and their specific needs.

How will we achieve it?

Our team has a strong technological expertise and has established a connection with local communities of the disabled people and with Varna Municipality. We have contacted and involved the local branch of the Union of the Blind People in Bulgaria, the local branch of the Union of Deaf People in Bulgaria and the local branch of the Union of Disabled People in Bulgaria. We hope that with their support and involvement we will be able to gather enough participants that can provide a stable focus group to form sustainable patterns and models.

Expecting results

Through gathering this information, the results will hopefully help us to fully understand the problems that different members of the community experience when moving through the city. This will help in making the necessary adjustments to the infrastructure and to develop innovative solutions that address their needs.

Moreover, all collected information will be published as OpenData so that it can be used by all companies and institutions that want to improve the lives of Varna's citizens.

We believe that every city should be as accessible as possible to all inhabitants.

4 Preliminary and exploratory activities for Varna Limitless Project

Legal study

What is GDPR?

At its core, GDPR is a new set of rules designed to give EU citizens more control over their personal data. It aims to simplify the regulatory environment for business so both citizens and businesses in the European Union can fully benefit from the digital economy.
The reforms are designed to reflect the world we're living in now, and bring laws and obligations - including those around personal data, privacy and consent - across Europe up to speed for the internet-connected age.
Fundamentally, almost every aspect of our lives revolves around data. From social media companies, to banks, retailers, and governments - almost every service we use involves the collection and analysis of our personal data. Your name, address, credit card number and more all collected, analysed and, perhaps most importantly, stored by organisations.
What is GDPR compliance?
Data breaches inevitably happen. Information gets lost, stolen or otherwise released into the hands of people who were never intended to see it - and those people often have malicious intent.
Under the terms of GDPR, not only do organisations have to ensure that personal data is gathered legally and under strict conditions, but those who collect and manage it are obliged to protect it from misuse and exploitation, as well as to respect the rights of data owners - or face penalties for not doing so.
What types of privacy data does the GDPR protect?

- Basic identity information such as name, address and ID numbers
- Web data such as location, IP address, cookie data and RFID tags
- Health and genetic data
- Biometric data
- Racial or ethnic data
- Political opinions
- Sexual orientation

Technical study

Varna Limitless project is an entire IT eco-system that contains several major parts:

- Mobile Apps
- Anonymization nodes
- Blockchain infrastructure
- Web-based interface for executing queries and building surveys
- Beacons
- IoT infrastructure (LoRaWAN)

On Figure 1 you can see a rudimentary model of different parts of the system and how it operates.
On Figure 1 you can see a rudimentary model of different parts of the system and how it operates. The client in this case NSI creates a campaing for data mining that trigger the creation of a smart-contract which will be responsible for the execution of the campaing and distribution of funds among the respondents. The smart-contract start the execution of the campaing by sending requests for

computations over their data to all respondents and sensors. Each one of them sends the result over the computation to the aggregation and anonymization nodes that collect and assemble all results. After that the accumulated results are sent back to the smart-contract and the client.

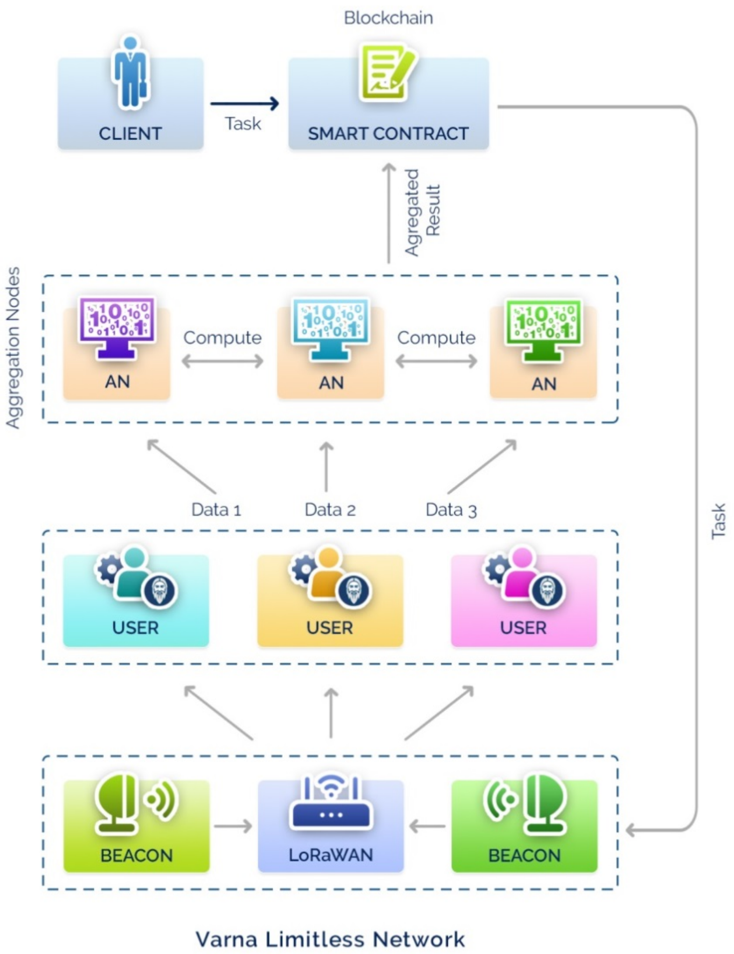

Figure 1 – Basic infrastructure of Varna Limitless

Our team studied the existing technologies that we need for building the project and the existing infrastructure in Varna that can be used. Our main focus was on anonymization modes and blockchain infrastructure because they are the most crucial part of the whole system. We want to provide respondents the ability to keep thier privacy and anonymity. This is the most vital part of Varna Limitless and the most unique one.

Mobile App

Our main source for gathering the movement data is the mobile APP that we developed. It is a standard Android APP, based on JAVA. Within the feasibility study, we were able to create some mockups that were discussed with the representatives of the Unions of disabled people.

On figures 2-7 you can see main screens from our mobile App. On figure 2 you can see the launch home screen where users have to input their password in order to proceed further into the app. On figure 3 is shown the main menu of Varna Limitless App. Users can from there navigate to sections Surveys, Profile, Help and Settings. On figure 4 is the screen where every user can select the type of disability from a drop-down options. On the next screen (figure 5) you can see the main list with all surveys and data-mining tasks that one user can participate in. Every user can choose to participate or not. On figure 6 is shown another screen from the profile section where you can reach either or basic profile info or the tracking info. On figure 7 you can see a history of your movement. It can only be reached by the user and none else.

Conclusion: There are no technical hurdles that can be a problem for the mobile APP development.

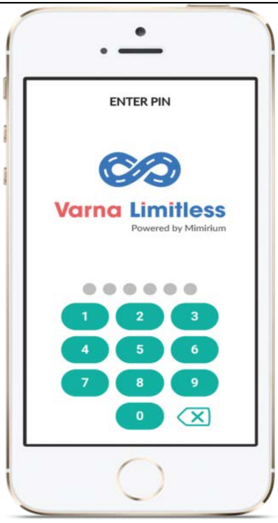

Figure 2 - Home screen mobile app

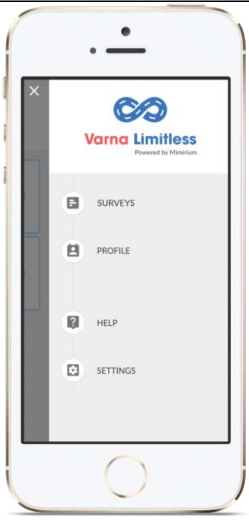

Figure 3 - Details

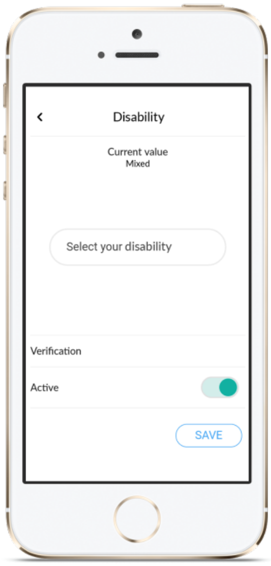

Figure 4 – Disability screen

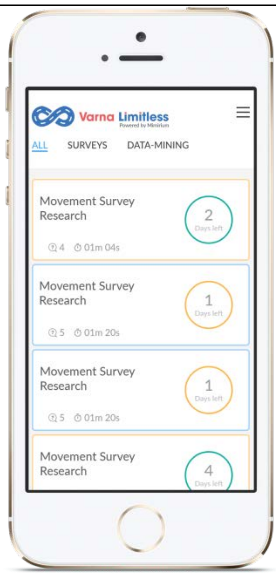

Figure 5 – Tasks screen

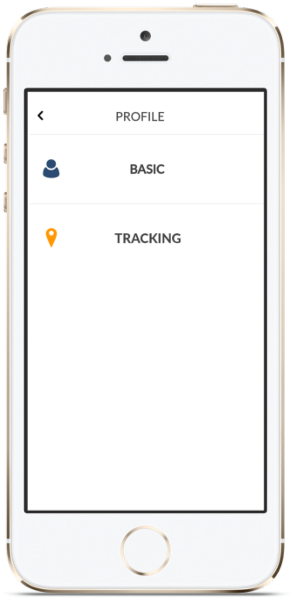

Figure 6 – Data collection screen

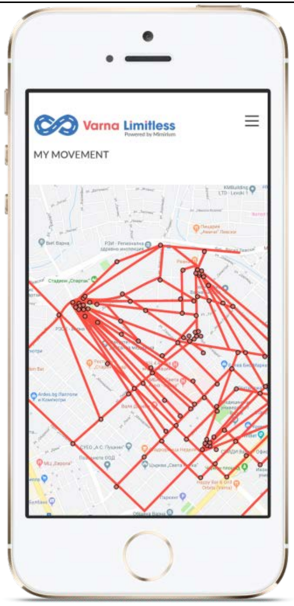

Figure 7 – My movement screen

Anonymization nodes

Our anonymization nodes[10] are the most essential part of the system. In order to work properly they have to utilize several cutting-edge technologies – **homomorphic cryptography, zero-knowledge proof and multipart computations**.

Homomorphic encryption is a form of encryption that allows specific types of computations to be executed on cipher texts and obtain an encrypted result that is the cipher text of the result of operations performed on the plain text. Applying the standard encryption methods induces a dilemma: If the data is stored unencrypted, it can reveal sensitive information to the storage/database service provider. On the other hand, if it is encrypted, it is impossible for the provider to operate on it. If data are encrypted, then answering even a simple counting query (for example, the number of records or files that contain a certain keyword) would typically require downloading and decrypting the entire database content. A homomorphic encryption allows a user to manipulate without needing to decrypt it first.

Zero-Knowledge Proofs (ZKPs)[11] allow data to be verified without revealing that data. They therefore have the potential to revolutionize the way data are collected, used and transacted with. Each transaction has a 'verifier' and a 'prover'. In a transaction using ZKPs, the prover attempts to prove something to the verifier without telling the verifier anything else about it. By providing the final output, the prover proves that they are able to compute something without revealing the input or the computational process. Meanwhile, the verifier only learns about the output.

A true ZKP needs to prove 3 criteria:

- *Completeness*: it should convince the verifier that the prover knows what they say they know.

[10] A node is a point of intersection/conncetion within a network. In an environment where all devices are accessible through the network, these devices are all considered nodes (see techopedia.com).
[11] More information about ZKPs - https://hackernoon.com/eli5-zero-knowledge-proof-78a276db9eff

- *Soundness*: if the information is false, it cannot convince the verifier that the prover's information is true.
- *Zero-knowledge-ness*: it should reveal nothing else to the verifier.

Multipart computation is a way that data can be used to analyse complex problems, shine a light on new solutions or even resolve otherwise unanswerable questions. But when it comes to using data for the public good, such as finding new drug targets for cancer or understanding how ride-sharing apps can influence traffic congestion, there is often societal tension between data sharing and data protection. In many cases, data sharing is even constrained or prevented by legal, ethical or privacy restrictions.

To harness the power of big data while maintaining privacy protection, researchers at Boston University's Rafik B. Hariri Institute for Computing and Computational Science & Engineering are leveraging a cryptography technology called secure multiparty computation (MPC), which allows collaborative data analysis without revealing private data in the process.

Through MPC protocol, parties enter their data, which is then split into separate pieces and masked with other random numbers; the encoded data pieces are sent to multiple servers, assuring data privacy. Organizations can, for example, input financial, personal, or patient data for comparison and analysis without ever receiving or seeing other parties' data.

Conclusion: All needed technologies are available and can be used in order to build the needed anonymization nodes.

DLT (Distributed Ledger Technology) infrastructure

The progress of humankind is marked by the rise of new technologies and the human ingenuity they unlock.

In DLT, we may be witnessing one of those potential explosions of creative potential that catalyse exceptional levels of innovation. The technology could prove to have the capacity to deliver a new kind of trust to a wide range of services. As we have seen open data revolutionise the citizens' relationship with the state, so the visibility in these technologies may reform our financial markets, supply chains, consumer and business-to-business services and publicly held registers.

There will be challenges as DLT matures and it will disrupt the way we generally think about and store data.

Algorithms that enable the creation of DLT are powerful, disruptive innovations that could transform the delivery of public and private services and enhance productivity through a wide range of applications.

Ledgers have been at the heart of commerce since ancient times and are used to record different things, most commonly assets such as money and property. They have moved from being recorded on clay tablets to papyrus, vellum and paper. However, in all this time the only notable innovation has been computerisation, which initially was simply a transfer from paper to bytes. Now, for the first time algorithms enable the collaborative creation of digital distributed ledgers with properties and capabilities that go far beyond traditional paper-based ledgers.

A distributed ledger is essentially an asset database that can be shared across a network of multiple sites, geographies or institutions. All participants within a network can have their own identical copy of the ledger. Any changes to the ledger are reflected in all copies in minutes or in some cases seconds. The assets can be financial, legal, physical or electronical. The security and accuracy of the assets stored in the ledger are maintained cryptographically through the use of 'keys' and signatures to control the usage within the shared ledger. Entries can also be updated by one, some or all of the participants, according to rules agreed by the network.

Distributed ledger technologies have the potential to help governments to collect taxes, deliver benefits, issue passports, record land registries, assure the supply chain of goods and generally ensure the integrity of government records and services. In the NHS, the technology offers the potential to improve health care by improving and authenticating the delivery of services and by sharing records securely according to exact rules. For the consumer of all of these services, the technology offers the potential, according to the circumstances, to control access to personal records and to know who has accessed them.

Existing methods of data management, especially of personal data, typically involve large legacy IT systems located within a single institution. To these are added an array of networking and messaging systems to communicate with the outside world, which adds cost and complexity. Highly centralised systems present a high cost single point of failure. They may be vulnerable to cyber-attack and the data is often out of sync, out of date or simply inaccurate.

In contrast, distributed ledgers are inherently harder to attack because instead of a single database, there are multiple shared copies of the same database, so a cyber-attack would have to attack all the copies simultaneously to be successful. The technology is also resistant to unauthorised change or malicious tampering, in that the participants in the network will immediately spot a change to one part of the ledger. Added to this, the methods by which information is secured and updated mean that participants can share data and be confident that all copies of the ledger at any time match each other.

The most popular form of DLT nowadays is blockchain. We will explore 3 of the most popular blockchains that support smart contracts.

Hyperledger[12]

Hyperledger Fabric is a blockchain framework implementation and one of the Hyperledger projects hosted by The Linux Foundation. Intended as a foundation for developing applications or solutions with a modular architecture, Hyperledger Fabric allows components, such as consensus and membership services, to be plug-and-play. Hyperledger Fabric leverages container technology to host smart contracts called 'chaincode' that comprise the application logic of the system. Hyperledger Fabric was initially contributed by Digital Asset and IBM, as a result of the first hackathon.

Hyperledger Burrow is one of the Hyperledger projects hosted by The Linux Foundation. Hyperledger Burrow provides a modular blockchain client with a permissioned smart contract interpreter partially developed to the specification of the Ethereum Virtual Machine (EVM).

Hyperledger Indy is a distributed ledger, purpose-built for decentralized identity. It provides tools, libraries, and reusable components for creating and using independent digital identities rooted on blockchains or other distributed ledgers so that they are interoperable across administrative domains, applications and any other 'silo'. Because distributed ledgers cannot be altered after the

[12] More information about Hyperledger - https://www.hyperledger.org/

fact, it is essential that use cases for ledger-based identity carefully consider foundational components, including performance, scale, trust model and privacy. In particular, Privacy by Design[13] (*'Privacy by Design' means nothing more than 'data protection through technology design'. Behind this is the thought that data protection in data processing procedures is best adhered to when it is already integrated in the technology when created*) and privacy-preserving technologies are critically important for a public identity ledger where correlation can take place on a global scale. For all these reasons, Hyperledger Indy has developed specifications, terminology, and design patterns for decentralized identity along with an implementation of these concepts that can be leveraged and consumed both inside and outside the Hyperledger Consortium.

Etherium[14]

Ethereum is the foundation for a new era of the internet:

- An internet where money and payments are built in.
- An internet where users can own their data, and your apps don't spy and steal from you.
- An internet where everyone has access to an open financial system.
- An internet built on neutral, open-access infrastructure, not controlled by a company or person.

Launched in 2015, Ethereum is the world's leading programmable blockchain. But unlike other blockchains, Ethereum can do much more. Ethereum is programmable, which means that developers can use it to build new kinds of applications. These decentralized applications (or 'dapps') gain the benefits of cryptocurrency and blockchain technology. They can be trustworthy, meaning that once they are 'uploaded' to Ethereum, they will always run as programmed. They can control digital assets in order to create new kinds of financial applications. They can be decentralized, meaning that no single entity or person controls them.

EOS[15]

EOS.IO is a blockchain protocol powered by the native cryptocurrency EOS. The protocol emulates most of the attributes of a real computer including hardware (CPU(s) & GPU(s) for processing, local/RAM memory, hard-disk storage) with the computing resources distributed equally among EOS cryptocurrency holders. EOSIO operates as a smart contract platform and decentralized operating system intended for the deployment of industrial-scale decentralized applications through a decentralized autonomous corporation model. The smart contract platform claims to eliminate transaction fees and also conduct millions of transactions per second.

The aim of the platform is to provide decentralized application hosting, smart contract capability and decentralized storage of enterprise solutions that solve the scalability issues of blockchains like Bitcoin and Ethereum, as well as eliminating all fees for users. EOSIO accomplishes this by being both multi-threaded (able to run on multiple computer cores) as well as using delegated proof-of-stake for its consensus protocol. It aims to be the first decentralized operating system (EOSIO) that provides a development environment for decentralized applications like Steemit, a social network with monetary incentives and BitShares, a decentralized cryptocurrency exchange (DEX).

---

[13] https://iapp.org/media/pdf/resource_center/Privacy%20by%20Design%20-%207%20Foundational%20Principles.pdf

[14] More information about Etherium - https://www.ethereum.org/

[15] More information about EOS - https://www.eos.io

5 Existing DLT infrastructure in Varna

There are 5 working nodes that run Hyperledger, Etherium and EOS especially for Varna based projects. They are distributed between five different entities (NGO Innovator, NGO VarnaLabs, NGO Innovative and Advanced Technologies, Varna University of Management and Mimirium Ltd.) in order to provide secure distribution and publicity of all data. All nodes are publicly open and everyone can use them to run a project for a Smart City project based in the area of Varna. There are plans to multiply the nodes and to add additional DLT services like Tron, Aeternity, IPFS etc.

Conclusion: The needed DLT infrastructure is available and can be used in order to build the needed smart-contracts.

Web-based interface for executing queries and building surveys

This is a fairly simple web-based project that will be developed in Angular 7.0, .NET Core 2.2 and MongoDB. All of these technologies are mature and are used in millions of projects all over the world.

Conclusion: No technical hurdles for developing the needed module. The estimated duration for this will be around 3 months.

Beacons

Technology can turn something we have come to known on its head without changing its core premise. Smartwatches are a good example. Sure, Pebble and other smart timepieces of the first wave might have looked like toys, but the latest products from Motorola and other tech companies can be confused for a real chronometer; they look like one, tell time in the same way, but also offer a range of other features backed up by app solutions.

Beacons are at least as old watches: eye-catching constructions that offer information and navigation on land and at sea (the most common example is probably a lighthouse). Nevertheless, the newly developed beacon technology is bringing new context to an old word. If today you ask somebody what a beacon is, chances are they will tell you something like this: it is a small device that allows you to create contextual experiences near or in an area where it’s installed.

Beacons are small computers, roughly the size of a standard Wi-Fi router. As part of indoor positioning systems, beacons use proximity technology to detect human presence nearby and trigger pre-set actions to deliver informational, contextual, and personalised experiences.

When a user walks past an area where an indoor positioning system is set up, a beacon sends a code with a message to their mobile device. Here app solutions come forth: this coded message, which is shown in a form of a notification, can only be viewed with a mobile app (third party or brand mobile app).

Conclusion: No technical hurdles for deploying and configuring the needed beacons.

IoT infrastructure - Long Range Wide Area Network (LoRaWAN)

LoRaWAN[16] is a media access control (MAC) protocol for wide area networks. It is designed to allow low-powered devices to communicate with internet-connected applications over long range

[16] More information about LoRaWAN - https://lora-alliance.org/about-lorawan

wireless connections. LoRaWAN can be mapped to the second and third layer of the OSI model. The Open Systems Interconnection model (OSI model) is a conceptual model that characterizes and standardizes the communication functions of a telecommunication or computing system without regard to its underlying internal structure and technology. Its goal is the interoperability of diverse communication systems with standard communication protocols. The model partitions a communication system into abstraction layers. The original version of the model had seven layers. It is implemented on top of LoRa or FSK modulation in industrial, scientific and medical (ISM) radio bands. Frequency-shift keying (FSK) is a frequency modulation scheme in which digital information is transmitted through discrete frequency changes of a carrier signal. The technology is used for communication systems such as telemetry, weather balloon radiosondes, caller ID, garage door openers, and low frequency radio transmission in the VLF and ELF bands.

Existing LoraWAN infrastructure in Varna

There are five gateways operating in Varna that cover approximately 75% of the city area. You can see their locations and approximate coverage on Figure 9.

The LoRaWAN Varna community represented by NGO VarnaLabs is planning to add 5 more gateways that will be enough to cover 100% of Varna area and provide reliable and secure backbone for IoT and smart device communication.

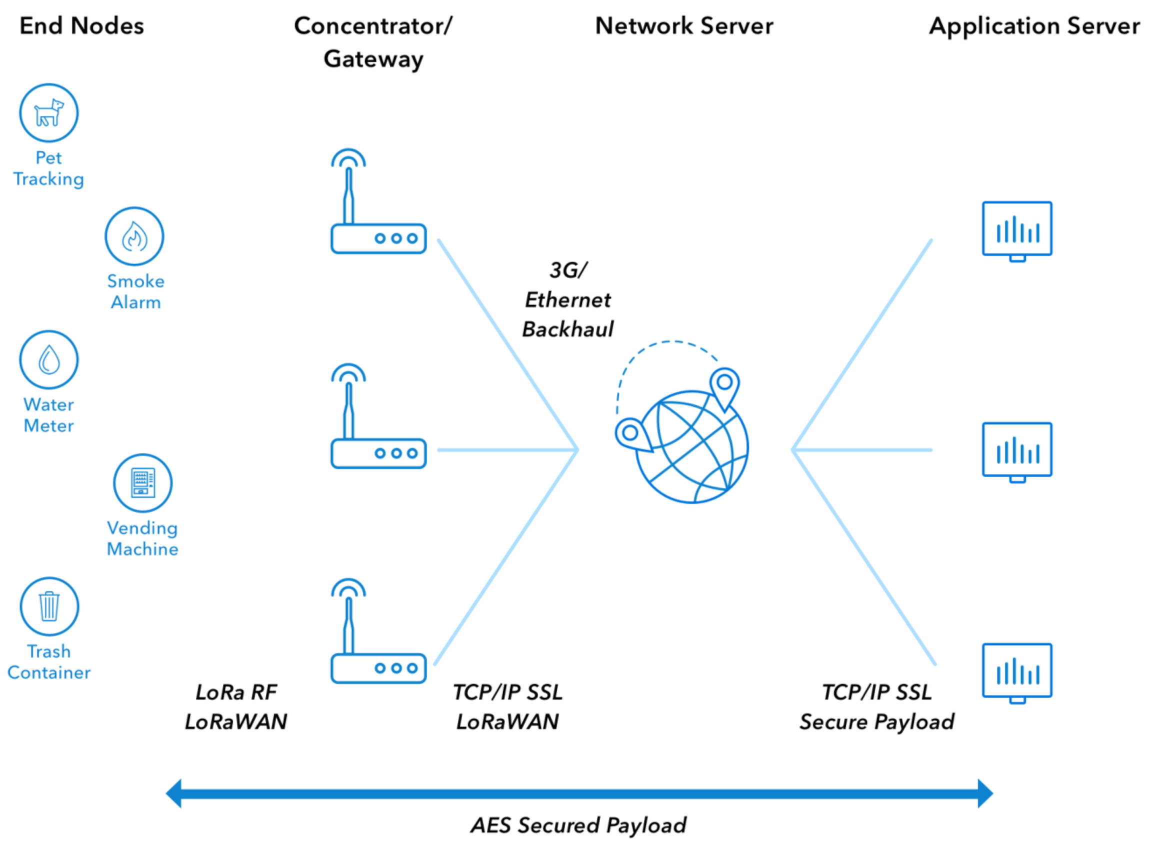

Figure 3 - LoRaWAN Infrastructure

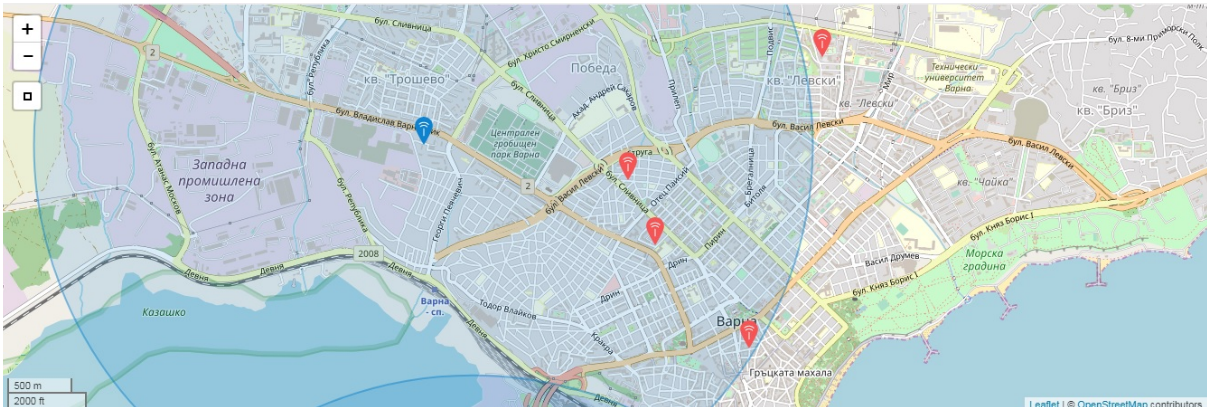

Figure 4 - LoRaWAN Gateways in Varna

Conclusion: The needed IoT infrastructure is available and can be used in order to connect the needed beacons.

Statistical and Social study

There is a tendency of an increasing number of people with disabilities both adults and children in Varna Municipality. Official data represented by Adriana Grigorova, Director of 'Social Assistance' Directorate - Varna show that there were almost 25,000 disabled people in Varna in the year 2018. There is an increase of about 4% compared to 2017. More than 1,000 children with permanent disabilities were assisted by the Directorate on various legal grounds last year.

Only last year, almost 57,000 Bulgarians with disabilities over 16 years have received an expert decision for permanently reduced working capacity for the first time. For 2018, more than 175,000 people with disabilities have been re-certified by expert medical committees or received their first decision. A vast majority of these people – more than 72% - are in retirement or retired. They are equally divided - 36% over the age of 60 and just as much in the age group of 50-59 years.

According to data from the Census 2011, there are 474 267 people with disabilities in Bulgaria; 9039 of them are children. During the Census 2011 there was a question of recognized permanently reduced working capacity or degree of disability.

Among the adults with permanently reduced working capacity, the share of people with disabilities is the largest - from 71 to 90% - 184,556. More than 90% of the injuries are 131,298 people. The lowest is the proportion of persons with reduced working capacity of 50% - 38,846. With increasing age, the share of people with up to 50% permanently reduced working capacity or degree of disability decreases.

Over 3030 of children have a disability of over 90%.

While children are dominated by disabled boys - 5203, women have the highest prevalence - 258,708, that is, every 1,000 women aged 16 and over, 112 have permanently reduced working capacity or disability, and for men - 96 per 1000.

About 2/3 of people with disabilities live in cities.

The highest is the number of people with permanently reduced working capacity or degree of disability in the age group 60-69 - 129 465 people.

Disabled people in Varna in the year 2018	25,000
Disabled people in Bulgaria in the year 2018	474,260

New certificates for permanently reduced working capacity for the first time 2018	57,000
Disabled children in Bulgaria in the year 2018	9,039

Conclusion: There are a lot of people in Bulgaria with different kinds of disabilities. Their number increase every year, because of the aging of the population and environmental factors. We need to address this issue by improving the infrastructure in order to provide a comfortable life and easy access to all city major parts. There is a huge potential of the scalability of the project not only in Bulgaria, but in the EU as a whole, because the statistics show that in average 17% of the EU citizens have different kinds of disabilities.

6 Execution of Varna Limitless Project

6.1 Participants

For the execution of our project, we selected two different groups of 10 individuals each. A group of 10 disabled people (all of them with partial loss of their movement capabilities with 25% to 50% degree of disability) and a control group of 10 people without disabilities. All of them males, between 35 and 40 years old, actively working, located in one particular district of the city (Levski) in order to keep the data consolidated. These restrictions were applied because for the purpose of this pilot project we need a compact group of individuals situated in a relatively small area of Varna city. By the data provided by The Union of Disabled People in Varna this group was the most suitable for this particular project (see Figure 10).

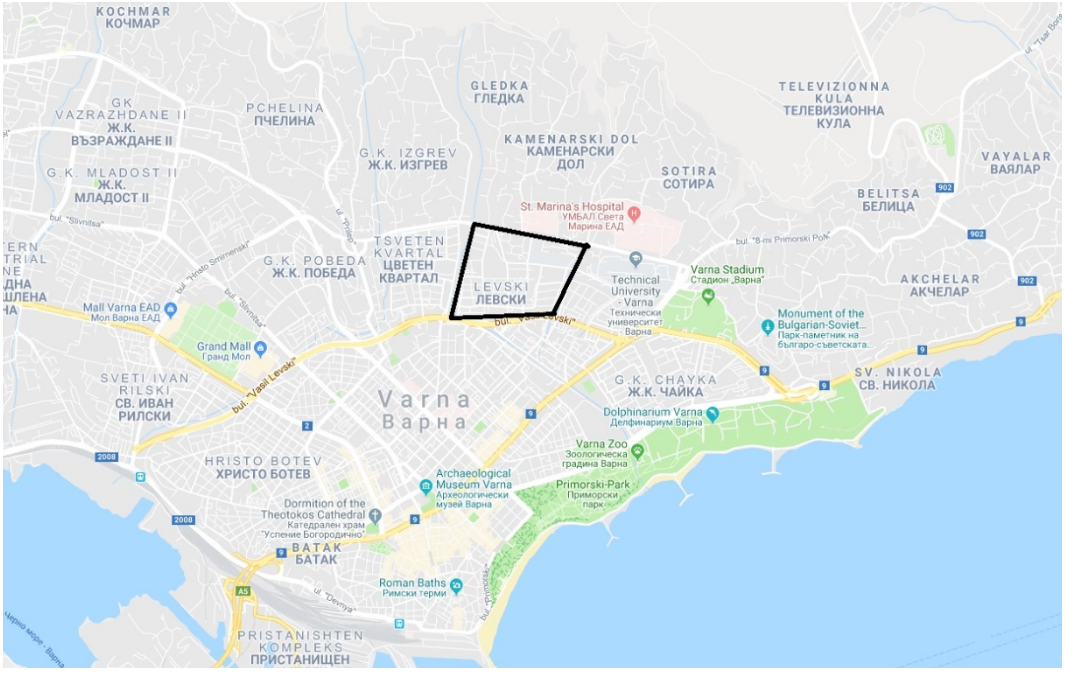

Figure 5 - Location area

6.2 Beacons

For this pilot project we set up a group of 20 beacons. 10 beacons were located in the living area of our participants and 10 beacons were located in different popular places in the city (Figure 11). The locations were:

In the living area:

- Jelezni vrata str. – bus stop
- Evlogi Georgiev str. – bus stop
- Evlogi Georgiev str. – University
- Podvis str. – bus stop
- Vasil Levski bul. – bus stop
- Mir str. – bus stop
- Dubrovnik str. – bus stop
- Dubrovnik str. – near Billa store
- Studentska str. – CBA store

- Studenstka str. – Elementary school

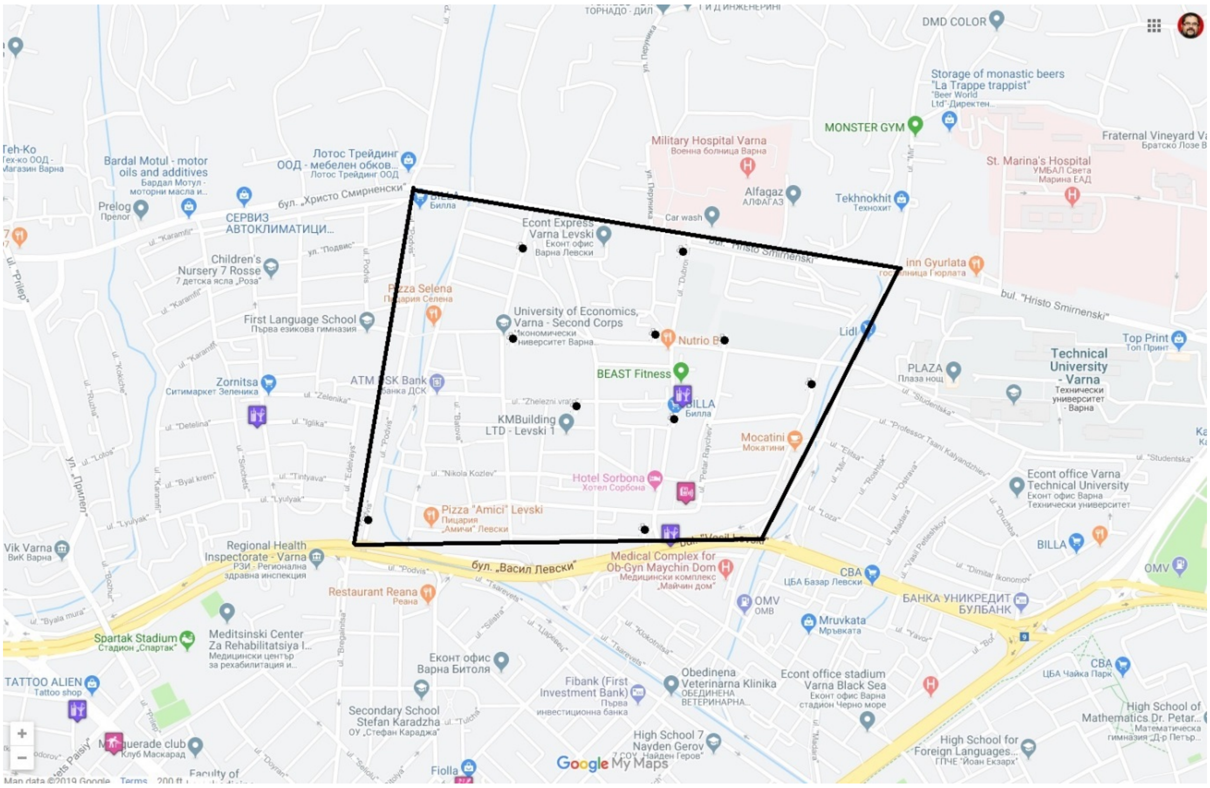

Figure 6 - Beacon locations in the living area

Other locations (Figure 7):

- Nezavisimost sqr.
- Sea Garden entrance
- Sevastopol sqr.
- Varna Municipality
- Dormition of the Theotokos Cathedral
- Collective farm market Varna
- The Red sqr.
- Palace of culture and sports Varna
- Tzar Osvoboditel bul. – Hospital
- Primorski bul. - recreational area

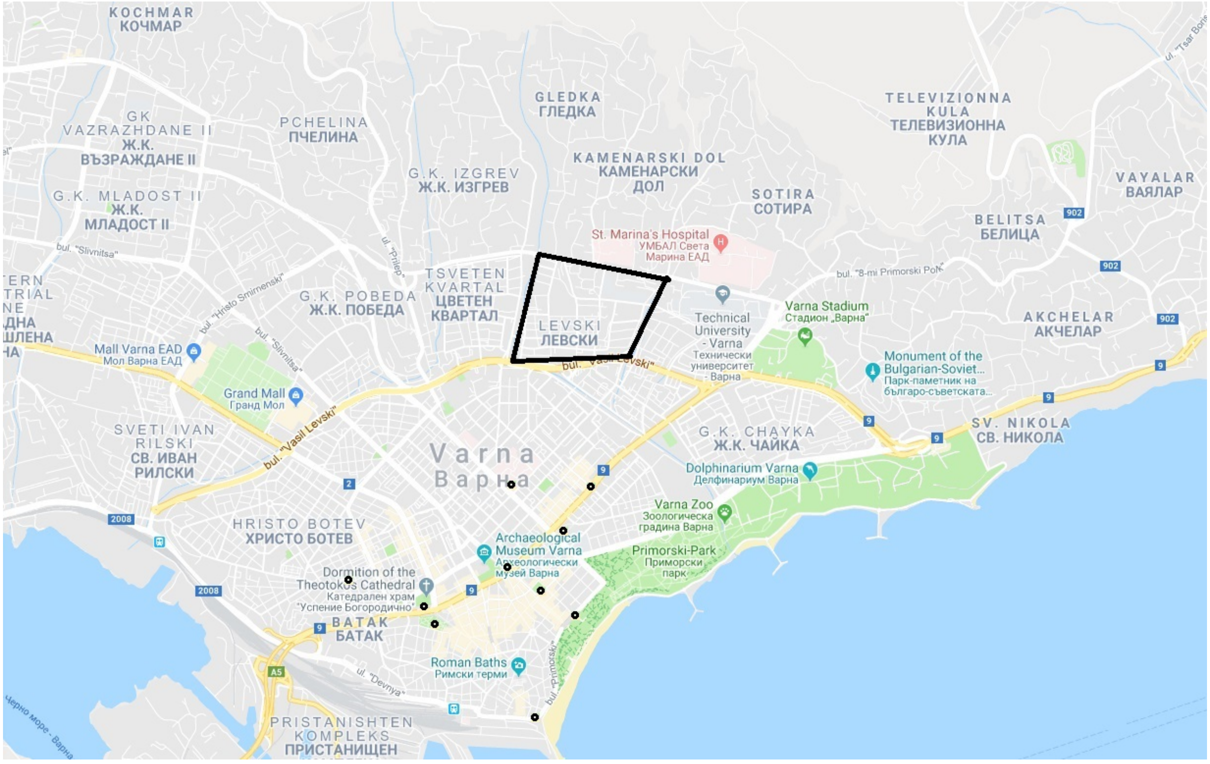

Figure 7 - Beacon locations outside the living area

All beacon locations were chosen based on previously collected data for the most visited places in Varna. Setting our beacons on these spots we were able to verify the validity, authenticity and reliability of the data gathered by our solution.

Varna Limitless system was launched in 1st of June 2019 and collected data till 31st of August 2019 as the first pilot set.

7 Results of Varna Limitless Project

At the start of the project we set to gather information for 5 different sets of data:

1. Common spots visited by citizens in Varna
2. Average number of visits to certain spots and places
3. Average speed of movement through the city of people without disabilities
4. Average speed of movement through the city of people with certain disabilities
5. Difference of the average speed between both groups

7.1 Common spots visited by citizens in Varna

In our 3 month pilot study our participants gathered information of their movement through the city saving their GPS data locally on their devices. At the end of the period, we were able to compare the gathered information without violating their privacy and extracting only anonymized and aggregated data. We get the following data for the most common places, visited by both groups. We exclude all places in the living area of the participants.

Place	Disabled people	Control group
1st most visited	Nezavisimost sqr.	Sevastopol sqr.
2nd most visited	Sevastopol sqr.	Nezavisimost sqr.
3rd most visited	Sea Garden entrance	Sea Garden entrance

7.2 Average number of visits to certain spots and places

Place	Disabled people	Control group
Nezavisimost sqr.	34	41
Sevastopol sqr.	32	37
Sea Garden entrance	28	34

7.3 Average speed of movement through the city

	Disabled people	Control group
Average speed of movement	2,9 km/h	4,2 km/h

7.4 Difference of the average speed of movement

During our pilot project, we were able to measure the difference between the average speed of movement between both groups. All speeds above 6,5 km/h and below 1 km/h were excluded in order to measure only the normal walking speed of both groups. The average difference is 1,3 km/h. Thanks to our mobile app we were also able to identify spots in the city where the average difference is greater than the usual. We pinpointed 3 different locations where the differences were as followed:

- Crossing of Roza str. and Vasil Levski bul. – 2,3 km/h
- A section of Podvis str. – 1,8 km/h
- A section of Dubrovnik str. – 1,7 km/h

We were able to identify the reasons of these differences and as it was planned it is because of problems with the infrastructure – in all of these places there are big constructions that make the movement harder and even much harder for disabled people.

8 Solutions for the production of official statistics based on Varna Limitless Project

Smart Statistics can be seen as the future extended role of official statistics in a world impregnated with smart technologies. Smart technologies involve real-time, automated, interactive technologies that optimize the physical operation of appliances and consumer devices. Statistics themselves would then be transformed into a smart technology embedded in smart systems that would transform ‘data’ to ‘information’.

Trusted Smart Statistics can be seen as a service provided by smart systems, embedding auditable and transparent data life-cycles, ensuring the validity and accuracy of the outputs, respecting data subjects' privacy and protecting confidentiality.

Mimirium with their pilot project Varna Limitless has proven that such kind of a smart system can be achieved and built even with the existing infrastructure and technologies. All collected data can be used as a base for official statistics and can give an interesting insight and even information that previously were not possible for collecting and verifying. The possibilities that distributed networks and solutions can give users are fascinating and we can achieve more trust and reliability. Based on such solutions we can minimize potential legal risks such as GDPR and E-privacy, we can minimize data breaches and data leakages. We can provide respondents 100% anonymity and security of their data while providing the NSI a powerful and flexible instrument for data mining of different types of data. Apart from data, based on the GPS information Mimirium app and infrastructure can also be used for:

- Gathering aggregated information from distributed network of IoT sensors
- Gathering aggregated and anonymized behavioral users’ data from mobile devices
- Gathering aggregated and anonymized behavioral users’ data based on their internet activities, including interests, purchases, searches etc.
- Gathering aggregated and anonymized behavioral users’ data based on their social network activities, cross-relating data from different sources

The possibilities are endless. In a more and more digitalized world solutions like Mimirium Network and Varna Limitless are just the beginning of a whole new era of user-driven and user-centric apps. Each time our future is more connected with information technology; therefore, we should be extremely careful on how we move forward in the internet world. We have to be sure that we can control our digital identities and all data we generate in the global network. We have to be sure that we receive a fair reward for our time and attention.

9 Recommendations

Existing methods of data management, especially of personal data, typically involve large legacy IT systems located within a single institution. To these are added an array of networking and messaging systems to communicate with the outside world, which adds cost and complexity. Highly centralised systems present a high cost single point of failure. They may be vulnerable to cyber-attack and the data are often out of sync, out of date or simply inaccurate.

In contrast, distributed ledgers are inherently harder to attack because instead of a single database, there are multiple shared copies of the same database, so a cyber-attack would have to attack all the copies simultaneously to be successful. The technology is also resistant to unauthorised change or malicious tampering, in that the participants in the network will immediately spot a change to one part of the ledger. Added to this, the methods by which information is secured and updated mean that participants can share data and be confident that all copies of the ledger match each other at any time.

Distributed ledgers offer a range of benefits to government and to other public and private sector organisations including statistics institutes. As their name implies, they can be distributed extremely widely in a precisely controlled fashion. They are highly efficient because changes by any participant with the necessary permission to modify the ledger are immediately reflected in all copies of the ledger. They can be equally robust in rejecting unauthorised changes, so corrupting the ledger is extremely difficult. However, distributed ledgers should not be seen as an end in themselves. It is only when they have other applications — such as smart contracts — layered on top on them, that their full potential can be realised.

<u>Based on the results of this pilot project we recommend:</u>

- NSIs should develop a high-level capability road map and a supporting outline plan based on the work of this report and other early stage activities that are already underway in departments
- We should invest further in the research required to ensure that distributed ledgers are scalable, secure and provide proof of correctness of their contents.
- We need to work with academia and industry to ensure that standards are set for the integrity, security and privacy of distributed ledgers and their contents.
- As well as top-down leadership and coordination, there is also a need to build capability and skills within NSIs. We recommend the establishment of a community of interest, bringing together the analytical and policy communities, to generate and develop potential 'use cases' and create a body of knowledge and expertise within the civil service.
- We should continue to explore the opportunities that Mimirium Network and Varna Limitless can provide us for collecting additional statistics data.

Case Study 2

1 Introduction

In case study 2, the three National Statistical Institutes (NSIs) of Italy (ISTAT) and Germany (Destatis and SSO BB) examined the exemplary cities Berlin and Darmstadt and 12 projects that are funded by the Horizon 2020 (H2020) programme of the European Commission to support smart cities and communities. The aim is to give a first overview of the different smart solutions that are partly still developed and partly already implemented in cities that are transforming to smart cities across Europe. These solutions focus explicitly on the three major domains energy, mobility and information and communication technologies (ICT).

Questions to be answered in this case study are:

1. Which information do cities and regions need in order to transform to smart cities and regions and which data could official statistics offer them?
2. Which data are generated by smart cities and regions and how could these data be used for the production of official statistics?

The European Commission's H2020 research and innovation programme has the aim of bringing together cities, industry and citizens to demonstrate solutions and business models that can be scaled up and replicated and that lead to measurable benefits in energy and resource efficiency, new markets and new jobs. These projects - called 'smart cities and communities lighthouse projects' - are facing the same challenges that other EU cities are experiencing regarding the expectations of citizens and enterprises in the era of IoT.[17]

Although there seems to be a lot of valuable data for official statistics, the availability and the access to the promising data sources is currently a huge challenge. This might be due to the fact that some projects are still ongoing, but also because some cities might not want to publish their data on a public website. The 'digital city' Darmstadt seems to be an exception in this matter with some real-time traffic data available online for free.

To answer the research questions of this case study, firstly, an overview of the main characteristics (duration, contacts, lighthouse cities serving as pilot cities for the implementation of the smart solutions etc.) of the H2020 projects was made. In a next step, the projects were categorised in three major domains: energy, mobility and ICT. The projects' websites and their publicly available documents and deliverables were consulted to gain a broader picture of the data that the projects can offer and that could be used for producing official statistics.

Therefore, chapter 2 gives a general overview of the Horizon 2020 projects. Chapter 3 highlights the findings regarding the German cities Darmstadt and Berlin. Chapter 4 summarizes the most relevant smart city projects and gives an overview about relevant topics. In chapter 5, information and data needs of the cities are described and in chapter 6 potential IoT data sources are identified that could be valuable for the production of official statistics. Chapter 7 includes an assessment of the current availability and accessibility of those potential data sources. Finally, chapter 8 tries to give some recommendations on the next steps to be taken and which topics could be analysed further.

A more detailed description of the 12 H2020 projects is listed in the appendix.

[17] https://ec.europa.eu/inea/en/horizon-2020/smart-cities-communities (accessed 19 September 2019)

2 General overview of the Horizon 2020 projects

	IRIS Smart Cities https://irissmartcities.eu	MAtchUP http://www.matchup-project.eu	Replicate https://replicate-project.eu/	STARDUST http://stardustproject.eu/
Duration	01.10.2017 - 30.09.2022	01.10.2017 - 30.09.2022	01.02.2016 -31.01.2021	01.10.2017 - 30.09.2021
Contact (Person)	Project coordinator: Haye Folkertsma Gemeente Utrecht (Contact Form)	Project coordinator: Ernesto Faubel Ayuntamiento de València efaubel@valencia.es Communication Secretariat: info@matchup-project.eu Costanza Caffo & Veronica Meneghello Fondazione iCons – iCube programme	Contact Form	Project coordinator: Florencio Manteca CENER – CIEMAT FOUNDATION fmanteca@cener.com Communication: Giulio Mazzolo communication@stardustproject.eu
Objectives	• Energy efficiency • Sustainable mobility • Integrated ICT infrastructures	• Energy efficiency in buildings an renewables • Sustainable mobility and logistic solutions • Integrated ICT infrastructure	• Saving energy consumption • Sustainable mobility • Integrated ICT infrastructures	• Saving energy and Reduce the greenhouse gas emissions by 63%. • Offer greener transportation • Creating and deploying open city information platforms
Light-house cities	1. Utrecht (Netherland) 2. Nice (France) 3. Gothenburg (Sweden)	1. Valencia (Spain) 2. Dresden (Germany) 3. Antalya (Turkey)	1. San Sebastián (Spain) 2. Florence (Italy) 3. Bristol (UK)	1. Pamplona (Spain) 2. Tampere (Finland) 3. Trento (Italy)
Follower cities	1. Vaasa (Finland) 2. Alexandroupolis (Greece) 3. Santa Cruz de Tenerife (Spain) 4. Focsani (Romania)	1. Ostend (Belgium) 2. Herzliya (Israel) 3. Skopje (FYROM) 4. Kerava (Finland)	1. Essen (Germany) 2. Lausanne (Switzerland) 3. Nilüfer (Turkey)	1. Cluj-Napoca (Romania) 2. Derry (UK) 3. Kozani (Greece) 4. Litoměřice (Czech Republic)
Link to project flyer	• IRIS Project Presentation • Info graphics	MAtchUP flyer	Replicate Dissemination Materials	STARDUST Leaflet

	GrowSmarter http://www.grow-smarter.eu/home/	SmartEnCity https://smartencity.eu/	Triangulum https://www.triangulum-project.eu/	Smarter Together https://www.smarter-together.eu/
Duration	01.01.2015 – 31.12.2019	01.02.2016 – 31.07.2021	02.2015 – 01.2020	01.02.2016 – 31.01.2021
Contact (Person)	Not available	Project Coordinators: Francisco Rodríguez Pérez-Curiel francisco.rodriguez@tecnalia.com Silvia Urra Uriarte silvia.urra@tecnalia.com TECNALIA Research & Innovation Phone: +34 946 430 850	Damian Wagner damian.wagner@iao.fraunhofer.de Fraunhofer IAO	Maxime Valentin mvalentin@lyon-confluence.fr SPL Lyon Confluence
Objectives	• Low Energy District • Sustainable Urban Mobility • Integrated Infrastructure	• Reducing energy demand and maximize renewable energy supply • Developing sustainable mobility • Intelligent use of ICT	• Reduction in energy consumption of buildings and shift to renewable energy sources • Increase electric vehicles and charging infrastructure • Dynamic ICT data hub (for smart energy/ mobility)	• Reduction in energy consumption of buildings and shift to renewable energy sources • Increase e-mobility • Data platform and monitoring systems
Light-house cities	1. Stockholm (Sweden) 2. Cologne (Germany) 3. Barcelona (Spain)	1. Vitoria-Gasteiz (Spain) 2. Tartu (Estonia) 3. Sonderborg (Denmark)	1. Manchester (UK) 2. Eindhoven (NL) 3. Stavanger (NO)	1. Munich (Germany) 2. Lyon (France) 3. Vienna (Austria)
Follower cities	1. Graz (Austria) 2. Porto (Portugal) 3. Suceava (Romania) 4. Cork (Ireland) 5. Valetta (Malta)	1. Lecce (Italy) 2. Asenovgrad (Bulgaria)	1. Leipzig (D) 2. Prague (CZ) 3. Sabadell (ES)	1. Santiago de Compostela (Spain) 2. Sofia (Bulgaria) 3. Venice (Italy)
Link to project flyer	Key GrowSmarter documents	SmartEnCity Flyer	Triangulum Brochure	

	REMOURBAN http://www.remourban.eu/	RUGGEDISED http://www.ruggedised.eu/	MY SMART LIFE https://www.mysmartlife.eu/mysmartlife/	Sharm-LLm/Sharing cities http://www.sharingcities.eu/
Duration	01.01.2015 – 31.12.2019	01.01.2016 - 31.12.2020	01.01.2016 - 31.12.2020	01.01.2016 – 31.12.2020
Contact (Person)	Project Responsible: Ángela Rivada arivada@ava.es Marketing Officer Contact: Ruth Stallwood Ruth.Stallwood@nottinghamcity.gov.uk Murat Aksu murat.aksu@tepebasi.bel.tr	General inquiries: info@ruggedised.eu	Rubén García Pajares CARTIF Technology Centre mysmartlife@mysmartlife.eu	Project Director: Nathan Pierce Bernadett Köteles-Degrendele Email: pmo@sharingcities.eu Communication, project coordinator: bernadett.degrendele@eurocities.eu
Objectives	• Improve energy efficiency • Transport sustainability and reducing emissions • Connect key information sources with city monitoring systems	• Energy reduction • Developing e-mobility solutions • Developing open data city platform	• Reducing CO2 emissions and increasing the use of renewable energy sources. • Involving citizens in the development of an integrated urban transformation strategy • Increasing digitalization through the urban platforms	• New energy efficient districts and shift to local renewable energy sources • Support green transport with electronic vehicles and infrastructure • Interactive representation of a sustainable model
Light-house cities	1. Nottingham (UK) 2. Valladolid (Spain) 3. Tepebasi/Eskisehir (Turkey)	1. Rotterdam (Netherland) 2. Glasgow (UK) 3. Umeå (Sweden)	1. Nantes (France) 2. Hamburg (Germany) 3. Helsinki (Finland)	1. Lisbon (Portugal) 2. London (UK) 3. Milan (Italy)
Follower cities	1. Seraing (Belgium) 2. Miskolc (Hungary)	1. Brno (Czech Republic) 2. Gdansk (Poland) 3. Parma (Italy)	1. Bydgoszcz (Poland) 2. Rijeka (Croatia) 3. Palencia (Spain)	1. Bordeaux (France) 2. Burgas (Bulgaria) 3. Warsaw (Poland)
Link to project flyer	About REMORUBAN REMOURBAN Brochure	About RUGGEDISED	My Smart Life flyer https://mysmartlife.eu/publications/	Leaflet http://nws.eurocities.eu/MediaShell/media/SharingCitiesLeaflet_all.pdf

3 Exploration of the German cities Darmstadt and Berlin

3.1 The 'digital city' Darmstadt

The German city Darmstadt has a strongly increasing population of 162,000 inhabitants and nearly 100,000 commuters per day into the city. Darmstadt offers approximately 134,000 jobs, has a well-known cluster of companies based in the chemistry, pharmacy, mechatronic and IT sector, 3 universities and a lot of scientific institutions, in the space and physics as well as the IT sector. According to a city ranking, which is produced annually by the magazine *WirtschaftsWoche* in partnership with the companies Immobilienscout24 and IW Consult, Darmstadt again took first place in the 'City of the Future' ranking in 2018. The ranking evaluates 71 independent German cities with a population of more than 100,000 people. Rankings are produced in three areas: level, dynamism and future. A wide range of factors are taken into consideration, including standard of living, the employment market, research performance, innovation and digitalization.

In 2017, Darmstadt has won a competition of the digital association Bitkom and now aims to be a living lab for testing smart city technologies in order to facilitate peoples' daily life in the city and to gain experience regarding the creation of long-term smart city solutions.

The city Darmstadt is a concrete example and its projects were explored regarding their usefulness for official statistics.

A first contact with the Statistics and Urban Research Department of Darmstadt was made at a Workshop at the end of January 2019. On February 26th and April 10th 2019 there were meetings with Statistics Darmstadt regarding the topic 'Smart City', but also regarding mobile network data. At both meetings, the private enterprise T-Systems, which is a subsidiary of 'Deutsche Telekom' and both a cooperation partner of Destatis and the city of Darmstadt, was attending as well.

At the first meeting in Darmstadt, Destatis presented the content of this case study and exchanged information with the representatives of Darmstadt and T-Systems regarding their involvement in the projects of the 'digital city' Darmstadt. The goal of the second meeting in Wiesbaden was to plan the next steps regarding a data map, which was supposed to give an overview of data sources and metadata of projects that are conducted within the context of the 'digital city' Darmstadt.

On May 7th, Destatis could attend a status and planning meeting of the 'digital city' Darmstadt. On June 4th, the representative of Statistics Darmstadt was attending the final meeting of WPL and presented some projects of the 'digital city' Darmstadt, also with regard to new digital data sources, their challenges, solutions and findings so far.

All the meetings and also some attempts via mail to get in touch with a representative of the 'digital city' Darmstadt that is responsible for the data platform and the executive manager, unfortunately did not lead to a significant progress in getting either access to more information about the projects that were already available online or to get some first data of the different projects. In order to build up a cooperation in this matter, more effort or rather another approach seems necessary.

The 'digital city' Darmstadt has three lighthouse topics:

1. Mobility and Environment: The goal is to become a digital pioneer in combining mobility and sustainability.
2. Digital Services and Society: The aim is to modernize administrative services and to enhance the digital competences of the citizens by education projects.
3. Economy and Technology: The goal is to provide a reliable and modern IT infrastructure for innovative companies, research and education institutes.

Besides, there are 14 areas of action; each contains several projects of different interest for official statistics. In the following list, areas of high relevance or interest for official statistics are listed first, whereas potentially less relevant areas will follow at the end:

1. Environment: In order to solve local environmental problems, data about the environment is needed for decision making and urban planning. The project 'smart waste' (3rd quarter 2018 – 4th quarter 2019) aims at the management of industry and household waste via sensors that measure the level of waste. Data are available for all involved in waste management. Therefore, the best collection time can be identified and empty runs are avoided. Furthermore, so-called 'CityTrees', biotech air filters that can be placed throughout the city, aim at controlling air pollution. The project of 'smart zoos' in which data such as temperature, air humidity, and feed level of animal enclosures is measured by sensors and transmitted by an app is aiming at work-relief and the improvement of work safety for animal caretakers.
   - ➔ Relevant for official statistics
2. Mobility: Because of an increasing traffic volume and therefore increasing demands on urban planning and environment protection, new mobility concepts are needed. The projects in this field are e.g. smart parking, which is the measurement of parking spaces via sensors and their communication via apps. Traffic and environment sensors (e.g. on traffic lights) are supposed to monitor urban traffic in real-time, data should then be available as open data. This includes the detection of vehicle types and speed to control and manage traffic. Further projects aim for free WiFi in buses and trams and the use of semi-autonomous trams and mini buses.
   - ➔ Of high relevance for official statistics; some open data is already available
3. Data Platform: The aim is to grant scientists, companies, but also citizens access to public data collected via sensors and infrastructure while protecting personal data. The project is planned to be finished by the 3rd quarter 2019.
   - ➔ High potential for official statistics
4. Health: In order to answer the increasing demand for medical care and to use resources and competencies efficiently, several projects were initiated. Besides the possibility to get a first medical consultation online, to check-in to hospitals online and to have a digital navigation in hospital buildings, it is planned to have a digital health platform on which the citizens of Darmstadt can manage all medical records easily themselves and that can also be used for a data exchange between medical institutions.
   - ➔ Digital health platform: interesting for official statistics, but high sensitivity of data
5. IT Infrastructure: The field of action includes a variety of projects. A broadband strategy aims at the extension of the public WiFi and the development of IT infrastructure for 5G and LoRaWAN (Long Range Wide Area Network). The idea of the project 'smart lighting' is that LEDs become brighter only when people or cars approach by using sensors with anticipatory lighting or sensors that count the stream of visitors. Also, a parking app is supposed to show the availability of parking spaces by using parking sensors. Pilots are scheduled until the end of 2019.
   - ➔ Smart lighting and especially the sensors that count the stream of visitors, could be relevant
6. Energy: The automatic control of energy consumption in public life and private households to pursue an efficient and low-emission energy management. Pilot projects concern intelligent street lighting and smart meters.
   - ➔ No further information available online, but generally of potential for official statistics
7. Public administration: The goal of this field of action is the digitalization of the public administration, e.g. online application management or e-files, in order to facilitate and accelerate

administrative procedures. Apparently it is planned to give scientists access to relevant data of the city. Unfortunately, it is not specified, which kind of data that could be.

➔ Might be relevant for official statistics

8. Cyber security: A high security of the above mentioned data platform and generally for the communication and storage of city data is aimed at.

   ➔ No further information available online, but probably not interesting for official statistics.
9. Education: This field of action includes pilot projects that aim at the exploration of digital teaching methods and the education of technical competences but also the reflection and critical awareness concerning digital devices and services.

   ➔ Not relevant for official statistics
10. Culture: The goal here is the examination of the possible connection between digitalisation and culture and the creation of an analogue platform that also enhances the dialogue in this matter.

    ➔ Not relevant for official statistics
11. Society: The projects in this field of action concern the participation of citizens in digital innovations and the digitalisation of political and societal processes.

    ➔ Not relevant for official statistics
12. Trade and Tourism: This area includes projects such as a digital shop window, which is supposed to be an online trading platform for local retailers with an integrated delivery logistic, and a mobile app for Darmstadt to provide information and services for tourists and residents.

    ➔ Not relevant for official statistics
13. Industry 4.0: The aim is to enhance the competitive position and the attractivity of the business location, e.g. by offering trainings and events for local businesses.

    ➔ Not relevant for official statistics
14. Security and civil protection: This includes the use of drones and body cameras for situation reports, but also a digital mobile application data entry for rescue services.

    ➔ Not relevant for official statistics

3.2 The ‘smart city’ Berlin

Since 1990 Berlin is one of the fastest growing cities in Germany. This applies both to an enormous population development and to a high economic growth especially in fields of culture, tourism and information technology. Related to an easy and unconventional, creative lifestyle, the city became a centre of attraction for young and highly qualified people both from Germany and from abroad and a large number of startups as well as innovative companies were established.

As a consequence, there is an increasing potential and great demand for smart technologies to improve the living conditions in Berlin in a multitude of fields, ranging from mobility over housing to a greener and better life in general. These changes require a modification of administration, politics, sufficient technical capacity and — of course — social networking and a network of experts to promote the ideas, drive the transformation and support the proponents to implementing their technical solutions. Regarding this, in 2015 the Berlin government adopted a ‘Smart City-Strategy’ in which different essential measures and challenges are listed. Apart from administration modernization, which, for instance, includes an open data platform and a consistent digital document management system for eGovernment, there are mentioned inter alia smart home, ambient assisted living, linkage to ICT, smart mobility and sustainable smart infrastructure. However, this strategy did not lead to very much activity yet.

The technology foundation Berlin (Technologiestiftung Berlin, short TSB) is one of the official key partners in carrying out projects to connect people in order to find new smart city solutions and bring them on a level of usability for the city. In an interview we received information about some important smart city projects in Berlin. Since there is no official monitoring of the smart city projects in place, in some cases it is rather difficult to report on the status of the ideas and solutions. Below, we compiled a list of some prominent projects which are on top of the official agenda, already implemented or—in our opinion—in some sense promising for the subject matter IoT or for official statistics.

- **District Energy Network**: There are several smaller individual solutions to manage energy consumption for higher efficiency and to use renewable energy in private and industrial buildings to reduce climate impact. The main purpose is to link these projects and to use the expertise to develop solutions that are operational on larger regional levels. Therefore, more data about state of renovation and energy sources as well as consumption are needed.
- **Smart Business District (InfraLab)**: To use synergies of different companies a cooperation between the administrations for water management (Berliner Wasserbetriebe*)* and waste diposal (Berliner Stadtreinigung) as well as the Berlin Public Transport (Berliner Verkehrsbetriebe) was formed. The aim of this collaboration is to plan a management system for rainwater usage, the production, distribution, storage and usage of electricity and heat energy as well as the creation of a common e-mobility infrastructure.
- **DataHub**: The government of Berlin intends to create a real time data hub to collect, aggregate, analyse and share data especially from IoT devices.
- **FixMyBerlin:** This Project was funded by the Federal Ministry of Transport and Digital Infrastructure to connect the administration and interested citizens to optimize urban planning and inform users of cycle ways about changes via interactive maps.
- **KiezRadar:** KiezRadar ensures citizen participation through an app to give information from local politics and administration about events as well as trends in their neighborhood. Since citizens can contribute in the development process of the app it will be a product created by the citizens.
- **TTN Network Berlin**: 'The Things Network Berlin' has set up a rather large LoRaWAN network already covering large areas of the city and still growing. This network serves as a rather slow but long range and free to use data network for several kinds of energy efficient IoT devices.
- **CityLAB**: A very important aspect in the whole subject matter of smart cities is the social one: It is necessary for the representatives from government, civil society, academia and start-ups to meet and to develop and think through new ideas to enhance the livability of the cities. The CityLAB is an experimental laboratory combining elements of a digital workshop, a co-working space and event space into a single location where participation and innovation are jointly pursued.
- **Badestellen:** A cooperation between administrations for water management (Berliner Wasserbetriebe*)* and State Office for Health and Social Affairs (Landesamt für Gesundheit und Soziales) and the use of open data for modeling processes results in an interactive map where citizens have the opportunity to get information about water quality of bathing places in Berlin.

Open data was a further aspect of our discussion with the TSB. Soon it was clear that most of the data needed for smart cities are data on a very regional level—e. g. data on the energy consumption of single households—which in a majority of cases cannot be provided by official statistics, but have to be obtained from regional companies or regional administrative data sources. Information privacy is a key matter here. Open data is used in a vast variety of smart city solutions and open data from official statistics is one important part of these, e.g. regional socio-demographic data and geocoded data on traffic accidents. We were told that the open data community (not only in Berlin) would appreciate a closer collaboration with official statistics.

4 Smart city projects – Overview about relevant topics

A rising number of urban challenges in cities have been observed. They concern air pollution, traffic congestion, fuel poverty, high energy consumption, a ‘digital gap’ and issues regarding economic growth and job creation for example. Therefore, areas such as energy, mobility, environment and ICT have been identified as key elements in improving citizens’ quality of life.
Some projects that follow the ‘Smart City’ approach - such as the twelve H2020 Projects and the exemplary chosen German smart cities Berlin and Darmstadt - have intelligent solutions for energy, mobility and ICT on their agenda, which will be integrated in cities together with innovative business models. Even the digitalization of the public administration, e.g. online application management or e-files, could contribute to improve citizens’ quality of life by accelerating administrative procedures. Generally speaking, with the objective of increasing the quality of life, smart cities are trying to reduce emissions via saving energy arrangements and boost e-mobility, raise the use of renewables and improve urban infrastructure while using smart devices in order to control and manage energy consumption and traffic. Smart solutions have been created to implement the smart city approach. Those solutions, which have shown their use in practice, shall prospectively serve as blueprints for their replication across Europe.

Section 1: Energy saving and integration of renewable energy sources
The cities of the H2020 projects as well as the German cities Darmstadt and Berlin are trying to save energy in order to improve buildings’ energy efficiency, partly even on district level (zero energy districts). For this, H2020 projects and Berlin are retrofitting, which includes for instance insulating building façade, the installation of multi-sourced heating like heat pumps, thermal storage and solar/photovoltaic plants. The integration of smart control systems designed to monitor and control the mechanical, electrical, lighting and other systems is planned. Renewable energy sources, smarter energy storage systems and lighting materials will be introduced to ensure and provide sufficient energy for the cities to make use of and to reduce CO_2 emissions. Smart grids as an energy network, which can automatically monitor energy flows, should document consumers and suppliers information on real-time consumption when coupled with smart metering systems. Darmstadt and Berlin pursued the approach of automatic control of energy consumption in private households (smart home) and in public life as well.

Section 2: Implementing sustainable e-mobility systems and improving urban infrastructure
For decreasing emissions, the deployment of e-mobility is planned. The cities want to provide electric vehicles, charging stations and e-bikes. They want to establish a smart sensor/smart point network to collect and return information from/to people and the urban environment. Different kinds of e-vehicles public sharing fleets are planned as well as smart street lighting with sensors. For example Darmstadt started projects like smart parking, which is the measurement of parking spaces via sensors and the implementation of traffic and environment sensors (e.g. on traffic lights) to monitor urban traffic in real-time.
Some cities of the H2020 projects have started projects to optimize waste management. The solution ‘smart waste’ of Darmstadt and some of the H2020-Cities aims at a more sustainable waste management, e.g. fit industry and household waste via sensors that measure the level of waste.
Also environmental conditions like CO_2, humidity, temperature, dust and noise will be tracked by traffic management systems. Darmstadt has installed so-called ‘city trees’ to control air pollution.

Section 3: Information and Communication Technology (ICT)
Furthermore, all smart city projects are planning to implement a kind of data platform/data hub, which collects the data of the used smart devices in the context of IoT. The aim is to monitor consumption ideally via real-time data, to analyze performance, to deliver support and services as well as driving digitalization and citizen participation. Here we might find concrete data sources (see chapter 3).
For all three sections, the usage of smart solutions is planned or already implemented. Here we find some starting points concerning identifying data needs of the smart cities (see chapter 2).

5 Data needs of smart cities

Identified data needs of the cities are as followed:
Section 1: Energy saving and integration of renewable energy sources

- Data of energy consumption of the city and region to compare and assess energy consumption in general.
- Data of energy statistics of the city/region, which shows the share of renewable power, e.g. solar/photovoltaic power, to assess the cities' position in producing renewable power.
- Data about cities and building, e.g. age of the buildings, age of power lines or cables, to find out, which residential area could usefully be converted into a zero energy district.
- Data and information about companies, which produce solar and photovoltaic plants that can help to pick out the best technology.
  Data and information, which is useful to identify areas suitable for the exploitation of geo thermal energy, e.g. area surface extension and its energetic potential, number of buildings that may be connected to the geo-thermal area etc.

Section 2: Implementing sustainable e-mobility systems and improving urban infrastructure

- Data about traffic to identify the most burdened routes.
- Data about the usage of public transport in comparison to car registration in certain selected areas to identify possible requirements for individual traffic.
- Data of household statistics to identify characteristics of single (small) residential areas to deduce their demands and needs from their ways of life on the infrastructure.
- Data and information about parking areas und power lines to identify parking spots for recharging e-vehicles.
- Data and information about the e-mobility endowment (infrastructures such as smart parking and service points; public, private and sharing e-vehicle fleets etc.) to evaluate the rank of the city in a green mobility potential perspective.
- Data about the amount of waste production in private households and commercial companies.

Section 3: Information and Communication Technology (ICT)

- Data about the used standards of digital data transfer respectively bandwidth for private households.

Concerning the example of the 'digital city' Darmstadt, there is an urban statistics department that conducts surveys if information is needed on a small-area level. Therefore, Statistics Darmstadt generally has data on commuters to Darmstadt, but even they lack on commuting data for small areas in Darmstadt. Mobile network data provides the needed small-area information about commuters in hourly intervals, e.g. the place of departure, the destination, such as tourist hot spots and big companies or the university e.g. Via the initiative 'digital city' Darmstadt and due to the won competition of the digital association 'Bitkom', Darmstadt is in a strong cooperation with T-Systems and got mobile network data for free in order to test its usability for their purposes.
Beside these suggestions, smart cities generally need data on a small-area level. That means ideally a level below the community/municipality level, e.g. on the district level. Therefore, official statistics might provide their data geocoded if possible, e.g. data of road traffic accidents.

Nonetheless small-area data on e.g. household level might be obtained from regional companies or regional registers, but questions of privacy and ownership must be clarified first.

6 Variety of promising data sources for official statistics

Smart solutions of the summarized areas of action above with relevance for the official statistics are listed below.

Section 1: Energy saving and integration of renewable energy sources

- Data of the smart power management systems, smart metering or smart control systems (real-time electricity consumption) could be relevant for statistics of energy consumption.
- Data on electricity consumption generated by smart devices on E-Vehicles could be relevant for statistics on emissions.
- Data of the solution ‘smart home’ could be collected and prepared for open data platform.
- Data of smart shell and equipment refurbishment is linked to open data platform.
- 3D-ICT tool to manage energy consumption in refurbishment process – database.

Section 2: Implementing sustainable e-mobility systems and improving urban infrastructure

- A ‘Smart Mobility Platform’ (gathering data from IoT) could be relevant for official statistics, e.g. traffic statistics, mobility statistics, road accidents statistics.
- Mobility behavior monitored by collecting data on an urban platform from a Multimodal hub where parking bikes and new pedestrian areas are coupled to public transportation and electrical vehicle chargers could be relevant for official statistics, e.g. mobility statistics
- Mobility behavior collected data from Cargo tricycles could be relevant for freight transport statistics.
- RFID (Radio-frequency identification)/Smart cards for multimodal transport could be relevant for official statistics on mobility behavior and passenger multimodality travelling.
- Data of implemented Apps of E-bike or E-cargo sharing models could be an addition for traffic statistics and freight transport statistics.
- Smart (LED) street lighting with sensors could be relevant for official statistics, e.g. traffic statistics, mobility statistics, road accidents statistics.
- Smart traffic signals/ traffic signal management system and traffic models could be relevant as addition to traffic statistics.
- Smart waste management integrating underground waste transportation, optical sorting technologies and data of level of accumulating waste could give data about waste handling.

Section 3: Urban data platform

Official statistics usually need a different kind of data, in particular less aggregated and more punctual data. Smart Platforms are surely the most important ‘solution’ for official statistics as well as they are designed to collect and store data in a standardized format, and at some extent to give back information that may be useful to build statistical indicators. Trusted Smart Statistics based on Smart Platforms may be an innovative challenge to produce information in the future; tools and apps that extract punctual (from a spatial and/or temporal point of view) data or even aggregated information to calculate indicators directly from Smart Platforms seem the most interesting way forward to exploit smart cities from a statistical point of view. In this perspective, the Smart Platforms presented in the H2020 projects and the cities Darmstadt and Berlin could be useful and valuable and enrich official statistics.

Examples for data platforms are:

- Monitoring district mobility patterns could be relevant for official statistics, e.g. for mobility or traffic statistics. (IRIS Smart Cities accessed 20.09.2019)
- Real-time data on the amount of solar power or geo-thermal power, stored in the storage systems in the districts, in Vehicle to Grid (V2G) e-car batteries[18], V2G e-bus batteries and stationary storage in homes and collective systems (IRIS Smart Cities accessed 20.09.2019) could be relevant for official statistics, e.g. statistics of electricity and heat generation of power plants for general care.
- Planned open city platforms with a holistic approach that will integrate all the urban elements from the sections ‘energy’ and ‘mobility’ could be a relevant data source for official statistics. (Stardust accessed 20.09.2019)
- Smart meters, urban and environmental sensors and Data Hub (Multiservice Concentrator; MSC)/ City Platform as an infrastructure of different utilities (electricity, water, heating). (GrowSmarter accessed 20.09.2019)
- Urban Cockpit analysing data of cities in real time (traffic, energy, environment e.g. CO_2, ambient light, air pressure, humidity, temperature, dust, noise) including a semantic model which reflects and connects the four domains energy, mobility, environment and ICT. (Grow Smarter accessed 20.09.2019)
- Urban management system, where monitored data from retrofitting packages and e-Vehicles are stored; communication with citizens. (SmartEnCity accessed 20.09.2019)
- Citizen engagement program; task to build IT platform for tracking energy consumption. (SmartEnCity accessed 20.09.2019)

[18] Vehicle to Grid (V2G) describes a system in which plug-in electric vehicles, e.g. such as battery electric vehicles communicate with the power grid to return electricity to the grid.

7 Data availability and accessibility

Based on currently available information from the evaluation of each single H2020 project, valid information about data structure, Key Performance Indicators and other figures are generally not available on most of the websites. This might be due to the fact that the projects are not closed yet. Only two projects will perspectively end by the end of 2019, most of them 2020 or later. More detailed analysis is necessary to figure out which nature and structure collected data of these different kinds of smart devices and smart meters have that are used in this city projects.
In particular, the analysis should focus on the data flow and from which data medium they will be transferred. So it becomes important and makes a difference if data are transmitted by direct access, 'pushing computation out techniques', by direct access to data producers platforms or access to citizen science platforms.
From the perspective of the production of official statistics it is important to keep in mind, that data collection via smart devices also means a huge number of users and high frequencies of data captured during the reference time. The technical facilities shall be in the position to handle this which should be considered in data analysis (aggregated data).
The accessibility of the data needs to be examined in more detail. Who owns the data? It is paramount to determine if data are of private, public or other kinds of property (PPP partners, community, etc.).

8 Recommendations

Among the large amount of proposed solutions, often similar in the various projects and in some cases pioneering or peculiar to a specific city framework, a few of them seem to respond more than others to the main characteristics that could enrich the production of official statistics.
In order to be more successful with getting in touch with smart cities representatives, we generally suggest the following actions:

- If possible, contact via already existing cooperation with private enterprises e.g. (in the case of Darmstadt this would be our contacts of the enterprise T-Systems)
- Invitations to interesting statistical conferences could also help catching their interest
- Writing a letter that is signed from the head of office. The letter should emphasize the benefit for the smart cities if they cooperate with us, e.g. hand-tailored analyses from the statistical offices for their purposes.

Use Case 1 – 'Digital city' Darmstadt:
The example of the 'digital city' Darmstadt seems a promising use case for further exploration, particularly as most of the projects will be finished at the end of the year 2019. Firstly, a collaboration with representatives of the city of Darmstadt and Statistics Darmstadt is already set up that could help for future work. Secondly, there is already some data available online and generally there are some promising areas of action for official statistics:

- Environment: especially the solution 'smart waste' that aims at the management of industry and household waste via sensors, e.g. to measure the level of waste
- Mobility: Smart parking – which is the measurement of parking spaces via sensors and their communication via apps – and smart traffic lights – that monitor traffic in real-time and detect vehicle types and speed
- Energy: Smart lighting – that is intelligent street lighting – and smart meters
- Data Platform: should contain public data collected via sensors and infrastructure

The first examination should concentrate on the data that is already freely available online and that is collected via the municipal traffic light systems. The data is real-time raw data from sensors on traffic lights that have detected increased traffic and that are updated every minute. This data can be used by citizens and researchers. The citizens can also find out about traffic restrictions in Darmstadt via a situation picture. The data is based on the sensor data combined with information about construction sites and other traffic restrictions.

The data deliver information about the location of the sensor, the occupancy rate (percentage of time a segment of the road is occupied by vehicles) of the detector and the number of cars detected in a minute.[19]

Use Case 2 – Data platform:
One main aspect of the smart city approach is the idea of collecting data of the used smart devices and smart sensors in a data platform. It seems to be the heart and central management system of the distributed devices and sensors, which are implemented in previously described areas of action

[19] https://darmstadt.ui-traffic.de/

like energy, mobility and urban infrastructure. As the projects described in this report are still ongoing, we don't know exactly how data is collected and which form and structure it has.
Promising projects for a follow-up study among the H2020 projects are especially GrowSmarter (until the end of 2019), SmartEnCity (until 2021) and IRIS Smart Cities (until 2022). They seem particularly promising, because the presentation of their projects in the internet is already quite detailed and with a special focus on their project plan concerning the data platform:

- GrowSmarter – Integrated infrastructures demonstrates different kinds of platforms (accessed 20.09.2019)
- SmartEnCity – Vitoria-Gasteiz and Sonderborg include an IoT platform for the integration of any sensor systems and like data acquisition layers, data models and data real-time repositories, security etc. (accessed 20.09.2019)
- IRIS – the listed Deliverables, especially about the planned City Innovation Platform, give a lot of insight to the platform architecture and the data management framework (accessed 20.09.2019)

A potential follow-up study should explore the fact in which extent the platforms are based on 'open' technologies and organized in a modular structure, so that they may be replicated, re-used and easily tailored to respond to the needs of the stakeholders in terms of information supplied (data flows, output, etc).

Use case 3 - Smart lamp posts:
As regards the choice of the 'very promising' solutions, in relation to the production of official statistics, we believe that the smart lamp posts could be very relevant in a possible future where streets will be equipped with them in a capillary way.
As they may be equipped with many sensors and additional devices, many experts agree to consider lamp posts as an innovative way to capture many different kind of data (vehicle traffic, 'active' traffic, air pollution, water level, and more).
In this context a possible use case could be, for example, in the field of road accidents statistics, where the big challenge is to estimate traffic flows punctually, in order to calculate precise indicators and indexes regarding the risk and probability of being involved in a road accident (the denominator for 'risk exposure' indicators). In this sense, the use of the lamp posts net to catch traffic flows may be used for this purpose.

9 Appendix

Horizon 2020 projects – objectives and solutions

9.1 IRIS Smart Cities

All Smart Solutions

Objective 1. *Saving energy at building/district level and maximum profits of renewables power/heat/gas*

- Positive Energy Buildings (solar PV; DC; heat pumps; batteries; heat storage) *[Gothenburg, Nice]*
- Near zero energy districts (solar PV; heat pumps; DC; refurbishment) [Utrecht, Nice, Gothenburg]
- Symbiotic waste heat (waste streams; circular economy; biofuels; energy management tools) [Nice]
- Multi-sourced district heating (excess heat from buildings equipment; heat pumps; thermal storage; decision and citizens apps) [Utrecht, Nice, Gothenburg]
- Local production of electricity with solar PV
- Share of locally produced and consumed renewable power at district scale
- Energy savings thanks to smart AC/DC power grid in apartments

Objective 2. *Local zero-emissions mobility and Lower household mobility costs*

- Participatory city modelling for energy and mobility (dynamic game map developed with Minecraft® involving citizenship), using data from sensors and apps [Gothenburg]
- Smart street lighting with devices [Utrecht]
- Solar powered V2G cars and solar/wind powered V2G buses [Utrecht]
- Solar-V2G charging stations/points [Utrecht, Nice, Gothenburg]
- E-vehicles public sharing fleets [Utrecht, Nice, Gothenburg]
- Geolocalization platform [Niece]
- EC2B mobility service (multimodal transport access system) [Gothenburg]
- Mobility behaviour monitoring [Utrecht, Nice]

Objective 3. *City Innovation Platform/Data Market*

- Urban monitoring services thanks to connected objects, i.e. smart sensors network (multi-sensoring) and the IoT deployed all over the city and to the Big Data processing capabilities. [Utrecht]
- A 3D CIM (City Innovation Model) that facilitates city management and planning by including building information, infrastructure, geodata and planning data. [Utrecht, Nice, Gothenburg]
- Through the City Innovation Platform:
    - Improve mobility services [Utrecht, Nice]
    - App for visualizing total energy use
    - Storage tracking at district level [Utrecht];
    - Early warning to Housing Corporation when households’ energy costs are rising [Utrecht];

- 'Energy Cloud' where near real-time data from energy (electricity, heat, water) consumption will be collected, integrated and made available for further analysis. [Utrecht, Nice, Gothenburg]

Smart Solutions (Devices/Systems) useful for official statistics

Objective 1. Saving energy at building/district level and maximum profits of renewables power/heat/gas

- Local production of electricity with solar PV
- Share of locally produced and consumed renewable power at district scale
- Energy savings thanks to smart AC/DC power grid in apartments

Objective 2. Local zero-emissions mobility and Lower household mobility costs

- Smart street lighting with sensors [purpose: *traffic statistics; mobility statistics; road accidents statistics*]
- E-vehicles public sharing fleets (data collection/management) [purpose: *mobility statistics*]
- Geolocalization platform [purpose: *traffic statistics; mobility statistics; road accidents statistics*]
- Mobility behaviour monitoring [purpose: *mobility statistics*]
- City Innovation Platform/Data Market [purpose: *traffic statistics; mobility statistics; road accidents statistics*]

Objective 3. City Innovation Platform/Data Market

- Monitoring district mobility patterns
- Real-time data on the amount of solar power stored in the storage systems in the district, in V2G e-car batteries, V2G e-bus batteries and stationary storage in homes and collective systems

Data Output

- Well defined list of KPIs (Report on the list of selected KPIs)
- City Innovation Platform (CIP)/Data Market

9.2 Match UP

All Smart solutions

Objective 1. Improvements in buildings’ energy efficiency and high integration of renewables in the energy supply

- Smart control systems designed to monitor and control the mechanical, electrical, lighting and other systems in a building, allowing the integration, automation, and optimization of any building system in support of facilities management and the building’s operation and performance
- Electrical storage: establish an optimized system at plant level that improves the peak regulation and increases the ability to accommodate renewable energy in the district heating networks, maximising energy savings, emission reduction and benefiting the regional power grid.
- Urban renewables: will be deployed at district and city level through the implementation of an innovative wave energy converter (WEC) system able to maximize the electric energy
- Smart grids as an energy network can automatically monitor energy flows and when coupled with smart metering systems, smart grids reach consumers and suppliers by providing information on real-time consumption. The project wants to implement smart power management systems, smart metering and smart interaction between energy networks.
- Public lighting: implementing urban micro-renewables, connectivity and smart control. In networked street lighting data and services are integrated.
- Electrical storage systems for self-consumption based on batteries, power supply system and charging stations will be integrated to improve the efficiency of the energy system at building level.

Objective 2. Increase the sustainable mobility and improve logistic solutions

- Smart street lighting with sensors [Valencia, Antalya]
- E-charging stations/points [Valencia, Antalya, Dresden]
- Electric or hybrid buses fleet; development and deployment of on-board equipment for electric buses (and related software) [Valencia, Antalya, Dresden]
- EVs for public fleet [Valencia, Antalya, Dresden]
- E-bikes [Antalya]
- E-bikes for disabled mobility [Valencia]
- E-bikes for last mile delivery [Valencia, Antalya]
- Multimodal hub/intermodal mobility point/modal integration [Valencia, Antalya, Dresden]
- Smart EV parking places (app) [Valencia]
- Mobility planning app [Dresden]
- Open Data Urban Platform [Valencia, Antalya, Dresden]

Objective 3. ICT Urban Platforms

- Implement specific IoT solutions linked to the platform and analyse the data and KPIs extracted from them to manage all of the city’s assets in the mobility, transport and energy sectors e.g. data from the vehicles chargers through the demand management systems. Electric vehicles will be monitored along the project and their data will be integrated into the urban platforms in order to determine their performance.

Smart solutions (Devices/Systems) useful for official statistics

Objective 1. Improvements in buildings' energy efficiency and high integration of renewables in the energy supply

- Data of the smart power management systems, smart metering or smart control systems

Objective 2. Increase the sustainable mobility and improve logistic solutions

- Smart street lighting with sensors [purpose: *traffic statistics; mobility statistics; road accidents statistics*]
- Multimodal hub/intermodal mobility point/modal integration [purpose: *mobility statistics; light utility vehicles (LUV) freight transport statistics*]
- Open Data Urban Platform [purpose: *traffic statistics; mobility statistics; road accidents statistics, light utility vehicles (LUV) freight transport statistics*]

Objective 3. ICT Urban Platforms, could be useful as the data source of all collected data

Data output

- Open Data Urban Platform
- data and KPIs extracted from specific IoT solutions
- The project creates an open data framework to foster the rise of new small businesses focusing on the provisioning of new added-value services in the ICT, mobility and energy sectors.

9.3 Replicate

All Smart Solutions

Objective 1. Saving energy consumption

- Building Retrofitting *[San Sebastian, Florence, Bristol]*
- District Heating (Renewable energy supply: biomass.) *[San Sebastian, Florence, Bristol]*
- Demand Side Platform (Monitoring of residents' consumption will be done fostering sustainable behaviour.) *[San Sebastian]*
- Smart grid and energy demand management *[Florence, Bristol*]
- Smart Info *[Florence]*
  - Device to monitor real-time electricity consumption
  - A supporting app, including heating consumption and mobility behaviours

Objective 2. Implementing sustainable e-mobility systems

- Clean Vehicles public fleet (e-/hybrid buses, motorbikes, bikes, taxis, cars) [San Sebastian, Florence, Bristol]
- Charging stations/points (public and private) [San Sebastian, Florence, Bristol]
- Smart street lighting with sensors [San Sebastian, Florence, Bristol]
- E-vans for delivery [San Sebastian]
- Advanced Mobility Service app (management of e-taxi service, bike sharing free flow and new additional e-car/e-van sharing) [Florence]
- Journey Planner Development app (It offers people in the Ashley, Easton and Lawrence Hill area an easy access to alternative and greener travel options) [Bristol]
- Parking app [Bristol]
- Smart Mobility Platform (gathering data from IoT; processing data to produce services of interest for cities and telephone companies) [San Sebastian]
- Smart City Platform [San Sebastian, Florence, Bristol]

Objective 3. Providing integrated ICT infrastructures that improve efficiencies in the use of local public resources and the delivery of Public Services

- Smart City Platform deployment with integrated services [San Sebastian, Florence, Bristol]
- Open Data and Citizen Participation services [San Sebastian, Florence, Bristol]
- High-speed connectivity network deployment for the whole city [San Sebastian]
- Smart lighting deployment [San Sebastian, Florence, Bristol]
- Smart IoT devices and the supporting infrastructure tested in urban environment [Florence, Bristol]
- Intelligent system: smart benches, smart waste bins, smart irrigation system, free Wi-Fi system, road car detection sensors, emissions level sensors [Florence]

Smart solutions (Devices/Systems) useful for official statistics

Objective 1. Saving energy consumption

- Renewable energy supply
- Real-time electricity consumption

Objective 2. Implementing sustainable e-mobility systems

- Smart street lighting with sensors [purpose: traffic statistics; mobility statistics; road accidents statistics]
- Advanced Mobility Service app [purpose: mobility statistics; light utility vehicles (LUV) freight transport statistics]
- Journey Planner Development app [purpose: mobility statistics]
- Smart Mobility Platform (gathering data from IoT) [purpose: traffic statistics; mobility statistics; road accidents statistics, light utility vehicles (LUV) freight transport statistics]

Objective 3. Providing integrated ICT infrastructures that improve efficiencies in the use of local public resources and the delivery of Public Services

- Maybe data from the intelligent system of smart benches, smart waste bins, smart irrigation system

Data output

- Smart City Platform - raw data/KPI list not available on project website/public deliverables (registration required).
- Mobility Related Services Provision: Mobility Diagnosis, O/D Matrix, Traffic Heat Maps, Modal Sharing Analysis, Reporting, etc.
- Big data and open data
- No list of KPI in the website

9.4 STARDUST

Relevant sub projects or goals

All Smart Solutions

Objective 1. Saving energy and reduce the greenhouse gas emissions by 63%.

- Introduce retrofitting and innovative heating and cooling systems to already existing buildings
- Developing efficient energy management protocols, user-designed interfaces of smart grids and storage systems and open sharing of data between users and other stakeholders. So, energy usage can be monitored and managed by the buildings' inhabitants and by the energy provider.
- Renewable energy sources and smarter energy storage systems and lighting materials will be introduced to ensure and provide sufficient energy for the cities to make use of.
- Increase citizens' quality of life through a smarter energy efficient system
- Installation of Active Roofs: Plug & Play;
- Biomass District heating in Txantrea, CHP with ORC technology;
- Microgrid in Ripagaina combined with smart lighting;
- Heat recovery in the municipal Data Centre;
- Heat recovery at urban scale: monitor and analysis of the potential use of water & waste water networks to integrate heat exchanger to provide heating & DHW through heat pumps;
- Smart lighting: A smart lighting control system, which takes additional factors to dimness and time of day into account;
- Lighting network to function as a platform for new smart services will be innovated
- 164 apartments for a total amount of 15,000 square meters: the realization of photovoltaic systems, a heat pump based on geothermal heat pumps, an innovative ventilated facade and an advanced system for monitoring consumption (applies to the the city Trento)

Objective 2. Offer greener transportation

- Providing electric vehicles, charging stations and other types of incentives that invite citizens to use
- E-taxis fleet/incentives (subsidies)
- Public E-car sharing
- EV purchasing incentives/free toll charges for EV
- E-bike sharing
- E-charging stations/points
- Solar taxi fast charging station.
- E-cargo bike last mile freight delivery
- Feasibility study of a last mile logistics centre with EV
- Smart sensors/smart points network to collect and return information from/to people and the urban environment

Objective 3. Data centres and infrastructures, and user-driven and demand-oriented city infrastructures will be introduced

- Smart sensors/smart points network advanced information points that will make the data collected by the sensors be usable [purpose: traffic statistics; mobility statistics; road accidents statistics]
- Installation of a sensor network to collect data on the environment, energy, mobility, safety and waste collection. The data will be processed and made available in real time to the administration and citizens through a dashboard and used to create advanced online services. [Trento]
- An Open City Platform with a holistic approach that will integrate all the urban elements from the section 'energy' and 'mobility' like for example smart lighting or eye-bike sharing system, E-charging infrastructure etc. [Pamplona]
- An open and common IoT platform, where all the data for example from different pilots and experiments will be brought, will be purchased for the city. [Tampere]

Smart solutions (Devices/Systems) useful for official statistics

Objective 1. Saving energy and Reduce the greenhouse gas emissions by 63%.

- Heat recovery at urban scale: monitor and analysis of the potential use of water & waste water networks to integrate heat exchanger to provide heating & DHW through heat pumps;
- A smart lighting control system

Objective 2. Offer greener transportation

- E-charging stations/points

Objective 3. Open City Platform

- ICT Urban Platforms, maybe useful as a data source of all collected (as a data hub) [purpose: traffic statistics; mobility statistics; road accidents statistics, light utility vehicles (LUV) freight transport statistics]

Data output

- Open City Platform - raw data/KPI list not available on project website/public deliverables.

9.5 GrowSmarter

All Smart Solutions

Objective 1. *Low Energy District*

- Climate shell refurbishment (insulate building facade, modern heat-pumps, self-regulating decentralised energy management system) *[Cologne]*
- Re-build of an industrial site: net zero energy buildings *[Barcelona]*
    - Photovoltaic (85 kWp)
    - Connection to district heating and cooling network
    - Energy management system
- Smart climate shell and equipment refurbishment (passive and active refurbishment) *[Barcelona]*
    - Apartment building (passive/ active measures)
    - Tertiary buildings (passive/ active measures, aerothermal heat pumps, photovoltaic)
    - Residential buildings (passive/ active measures, district heating,
- Energy efficient refurbishment of residential building *[Stockholm]*
- Passive and active energy refurbishment of swimming pools *[Barcelona]*
- Construction Consolidation Centre to improve production process
- Smart Home for tenants to save electricity and warm water (RheinEnergie) *[Cologne]*
- Active House smart home for tenants to save energy in the fields of electricity, hot water and heating *[Stockholm]*
- Open home Net system gives information about energy performance and a platform for all technical systems *[Stockholm]*
- Energy Saving Centre (Energy Management System) *[Stockholm]*
- Home Energy Management System *[Barcelona]*
- Virtual Energy Advisor analysing data from different devices *[Barcelona]*
- Stochastic model of appliances energy consumption *[Barcelona]*
- Virtual power plant (Siedlungsmanagement) which connects local photovoltaic production, heat pumps and batteries *[Cologne]*
- smart energy and self-sufficient block (photovoltaic and storage batteries / Energy Management *[Barcelona]*
- Building Energy Management System to monitor consumption of fossil fuels and electricity *[Barcelona]*

Objective 2. *Sustainable Urban Mobility*

- Smart traffic signals/ traffic signal management system and traffic models [*Stockholm, Barcelona*]
- Charging stations for car sharing companies and other customers *[Cologne]*
- Public fast charging stations *[Stockholm, Barcelona]*
- Alternative renewable fuelling stations *[Stockholm]*
- Charging stations for vehicle to building (V2B) and vehicle to grid (V2G/ V2X) application, Energy Management System *[Stockholm]*
- Smart taxi stand system (taxi parking sensors) *[Barcelona]*

- Car-sharing and Mobility ticket as a multimodal card for public and shared mobility *[Cologne]*
- Urban consolidation centre (UCC) for micro delivery of goods *[Barcelona]*
- Public delivery boxes and cargo bicycle deliveries *[Stockholm]*

Objective 3. *Integrated Infrastructures*

- Smart outdoor lighting with modern control system to save energy and costs (standalone system with automatic dimming curves, adjustable dimming curve in CMS, presence detecting system raising lighting levels) *[Stockholm]*
- Smart multifunctional tower for hyper connected space in cities and for monitoring the city *[Barcelona]*
- Smart Connected Street Environment monitoring transport and emission *[Stockholm]*
- Smart meters, urban and environmental sensors and Data Hub as an infrastructure of different utilities *[Barcelona]*
- Open District Heating using waste heat from cooling system for district heating network *[Stockholm]*
- District heating and cooling rings *[Barcelona]*
- Smart local thermal grids *[Barcelona]*
- Smart waste management integrating underground waste transportation and optical sorting technologies *[Stockholm]*

Smart Solutions (Devices/Systems) useful for official statistics

Objective 1. *Low Energy District*

- Climate shell refurbishment (insulate building facade, modern heat-pumps, self-regulating decentralised energy management system) *[Cologne]*
- Re-build of an industrial site: net zero energy buildings (Barcelona) integrated with measure 8.1 ‘Big consolidated open data platform’ and measure 8.2 ‘Urban models’ (open data platform/ urban model) *[Barcelona]*
- Smart Home (RheinEnergie) --> data will be collected and prepared for open data platform *[Cologne]*
- Self-regulating decentralised energy management system which is monitoring the electricity and heat production (connected with ‘virtual power plant’ from RheinEnergie) *[Cologne]*
- Smart shell and equipment refurbishment is linked to open data platform and urban model *[Barcelona]*

Objective 2. *Sustainable Urban Mobility*

- Smart traffic signals/ traffic signal management system and traffic models [*Stockholm, Barcelona*]
- Charging stations for car sharing companies and other customers *[Cologne]*
- Charging stations for vehicle to building (V2B) and vehicle to grid (V2G/ V2X) application, Energy Management System *[Stockholm]*
- Smart taxi stand system (taxi parking sensors) *[Barcelona]*

- Car-sharing and Mobility tickets as a multimodal card for public and shared mobility *[Cologne]*
- Cargo tricycles equipped with sensors *[Cologne; purpose: light utility vehicles (LUV) freight transport statistics*]
- Urban consolidation centre (UCC) for micro delivery of goods *[Barcelona]*
- Parking app *[Cologne; purpose: mobility statistics]*

Objective 3. *Integrated Infrastructures*

- Smart multifunctional tower for hyper connected space in cities and for monitoring the city *[Barcelona]*
- Smart Connected Street Environment monitoring transport and emission (Big Data Platform collecting people and vehicle flow - IBM)*[Stockholm]*
- Smart meters, urban and environmental sensors and Data Hub (Multiservice Concentrator; MSC - Endesa)/ City Platform as an infrastructure of different utilities (electricity, water, heating) *[Barcelona]*
- Smart local thermal grids *[Barcelona]*
- Urban Cockpit analysing data of cities in real-time (traffic, energy, environment) *[Cologne]*
- Urban Traffic as a fast and easy overview of current traffic flow in Cologne *[Cologne]*
- Big Open Data Platform including a semantic model which reflects and connects the three domains energy, mobility and ICT - *[Cologne]*
- Urban Environment gives overview about environmental condition (CO_2, ambient light, air pressure, humidity, temperature, dust, noise) provided by traffic management system - *[Cologne]*

Data output

- Derived from the Data management plan
    - Raw data from monitoring the electric, gas, and hot water consumption in buildings
    - Energy data from smart plugs distributed in households
    - Raw data from monitoring the electric, gas, and hot water consumption in residential and tertiary buildings. Indoor and outdoor temperatures, humidity, and solar radiation
    - Electricity production of the photovoltaic installation, client consumption
    - Heating: Raw data from waste heat production from data centre and supermarket
- Car Sharing: Data of car using, emissions and consumption of energy
- Parking app: Data of parking bookings and emissions
- Bike Sharing: Data of bike using, and consumption of electricity by e-bikes
- Measure 9.2 – Raw data on location and other parameters of electric cargo tricycles
- Measure 11.2 – Data of fast-charging events.
- Measure 12.5 – Motor sharing data on trips performed
- Measure 12.6 - Raw data obtained from sensors on occupation of taxi parking stands.
- Raw data related to charging of vehicles
- Electrical and cargo car pool: Raw data of users

See deliverable for more: http://www.grow-smarter.eu/fileadmin/editor-upload/Reports/D1.2_Revised_data_management_plan_2017-04-27.pdf

- Data of parking bookings and emissions
- Traffic lights, traffic *(also relevant for smart mobility?)*
- Raw data from electricity use for street lighting
- Raw data from waste handling
- Sensor data from street environment *(also relevant for smart mobility?)*

9.6 SmartEnCity

All Smart Solutions

Objective 1. *Smart Energy*

- Vitoria-Gasteiz retrofitting package *[Vitoria-Gasteiz]*
- Biomass district heating system with a resource management centre to optimize heating production and distribution depending on energy consumption (monitoring system) *[Vitoria-Gasteiz]*
- Tartu retrofitting package in addition with smart home system *[Tartu]*
- Smart home solution [Tartu]
- District cooling system that uses residual heat and PV panels *[Tartu]*
- Reusing old EV batteries with use of PV panels *[Tartu]*
- Retrofitting package including insulation, LED street lighting automated heating control system of mixed heating sources and PV plants *[Sonderborg]*

Objective 2. *Smart Urban Mobility*

- Public E-bike sharing system (use of a city Smart Card or mobile app [*Tartu*]
- Gas buses in the whole city (see also 'Smart Infrastructure') *[Tartu]*
- Acquisition of EVs (taxis and private cars) [*Vitoria-Gasteiz*]
- New biogas buses and biogas filling stations *[Sonderborg]*
- Intelligent EV chargers that charge, when lowest prices *[Sonderborg]*]

Objective 3. *ICT urban platform*

- Urban management system [Vitoria-Gasteiz]
- Citizen engagement strategy for the retrofitting package [Vitoria-Gasteiz]
- LED lights with smart controllers and light, noise and environmental sensors as well as cameras for detecting people and vehicle flow *[Tartu]*
- Citizen engagement program *[Sonderborg]*

Smart Solutions (Devices/Systems) useful for Official Statistics

Objective 1. *Smart Energy*

- Biomass district heating system with a resource management centre to optimize heating production and distribution depending on energy consumption (monitoring system) *[Vitoria-Gasteiz]*
- Tartu retrofitting package in addition with smart home system *[Tartu]*
- Smart home solution [Tartu]
- Retrofitting package including insulation, LED street lighting outmated heating control system of mixed heating sources and PV plants *[Sonderborg]*

Objective 2. *Smart Urban Mobility*

- Public E-bike sharing system (use of a city Smart Card or mobile app); collecting information about performance and use of the system [*Tartu*]
- New biogas buses and biogas filling stations; digital services to inform users of departures, arrivals, delays *[Sonderborg]*
- Intelligent EV chargers that charge, when lowest prices *[Sondergorg]*]

Objective 3. *ICT urban platform*

- Urban management system where monitored data from retrofitting package and e-vehicles are stored; communication with citizens *[Vitoria-Gasteiz]*
- LED lights with smart controllers and light, noise and environmental sensors as well as cameras for detecting people and vehicle flow *[Tartu]*
- Citizen Engagement - Technical consultations and community meetings *[Tartu]*
- Citizen engagement program; task to build IT platform for tracking energy consumption *[Sonderborg]*

Data output

Compiled from the project's data set description , ch. 6.1.2

- Quantitative (numerical) data that is derived using automated systems that allow collecting and store continuous flow of data.
    - Data registered by sensors and monitoring equipment (energy, electricity, temperature, etc.)
    - Data from the ICT system/platform that automatically stores variables that have been defined previously, i.e. for ICT and mobility
- City Information Open Platform (CIOP) - raw data/KPI list not available on project website/public deliverables. https://smartencity.eu/media/smartencity_d6.3_data_model_architecture_implementation_v1.0.pdf

9.7 Triangulum

All Smart Solutions

https://www.triangulum-project.eu/wp-content/uploads/2018/10/2018-01_D6.2-Smart-City-Framework.pdf

Objective 1. *Smart Energy*

- Central renewable energy plant (CEP); heats and cools with energy from city's waste and rain water [Stavanger]
- Building Energy Management System (BEMS) in Art Gallery *[Manchester]*
- 3D-ICT tool to manage energy consumption in refurbishment process *[Eindhoven]*
- Smart Gateway for homes [Stavanger]
- Smart Gateway for nursing homes [Stavanger]
- Smart Gateway for schools [Stavanger]
- Sewage heat pump system [Stavanger]
- Wind energy for common areas of apartment building (energy statistics) *[Eindhoven]*
- Micro-grid management system *[Manchester]*
- Demand Side Response Control for Student Accommodation *[Manchester]*
- Demand Side Response Control for Office Block (Academic Building) *[Manchester]*
- Demand Side Response Control for Public buildings *[Manchester]*

Objective 2. *Smart Mobility*

- Electric battery buses [*Stavanger*]
- Electric-Assist Cargo Bike hire [*Manchester*]
- Precise parking management information system [*Eindhoven*]
- Last mile delivery cargo bikes [*Manchester*]

Objective 3. *ICT*

- Cloud Data Hub (ICT) gathering and analysing data from 23 partners – access for Smart City Projects *[Stavanger]*
- Data curation, visualization and utilization (www.manchester-i.com) (ICT) *[Manchester]*
- Route of smart LED dimmable lights (ICT) *[Eindhoven]*
- Smart street lighting system including app for parking management and charging stations - backbone [*Eindhoven*]
- Blink: Innovative video for distance health care *[Stavanger]*
- Blink: Innovative video for communication services *[Stavanger]*
- Data Analytics Toolkit *[Stavanger]*
- Multimodal decision support service *[Stavanger]*
- Information hub for schools *[Stavanger]*
- Computing Platform *[Stavanger]*
- Data-Enabled Innovation Challenges [*Manchester*]
- App to train electric vehicle drivers [*Manchester*]
- Eindhoven Open Data Portal [*Eindhoven*]
- Public Sound Sensor Safety Project in Statumseind [*Eindhoven*]

Smart Solutions (Devices/Systems) useful for Official Statistics

Objective 1. *Smart Energy*

- 3D-ICT tool to manage energy consumption in refurbishment process - database *[Eindhoven]*
- Wind energy for common areas of apartment building (energy statistics) *[Eindhoven]*

Objective 2. *Smart Mobility*

- Electric battery buses; collecting data about CO_2 reduction possibilities [*Stavanger*]
- Electric-Assist Cargo Bike hire [*Manchester*]

Objective 3. *ICT*

- Cloud Data Hub (ICT) gathering and analysing data from 23 partners – access for Smart City Projects *[Stavanger]*
- Data curation, visualization and utilization (www.manchester-i.com) (ICT) *[Manchester]*
- Data Analytics Toolkit *[Stavanger]*
- Multimodal decision support service *[Stavanger]*

Data output

- Cloud Data Hub for Big Data (University of Stavanger)
- Smart city open dataset
- Cloud Data Hub: https://www.triangulum-project.eu/wp-content/uploads/2019/03/2018-01_D2.2-Cloud-Data-Hub.pdf :
- Data Management Plan http://www.triangulum-project.eu/wp-content/uploads/2017/12/D1.8_Revised_Data_Management_Plan.pdf
- *KPIs:*
    - Modal split (passengers/freight)
    - Carbon emissions
    - Average journey time
    - Average journey cost
    - Use of E-bike/E-car rental (GIS tracking)
    - Number of daily deliveries
    - Average delivery cost
    - Public transport use
    - Bicycle use
    - Use of E-buses (GIS tracking)
    - See deliverable for more: http://www.triangulum-project.eu/wp-content/uploads/2017/12/D2.1_Common_Monitoring_and_Impact_Assessment_Framework__V2.0_.pdf

9.8 Smarter Together

All Smart Solutions

Objective 1. *Smart Energy*

- Experiment with low energy districts providing energy-efficient buildings with local renewable heat and electricity
- 151.800 m2 photovoltaic systems (1MWp) and a wood-fired co-generation power plant
- Electric-Renewable Energy Sources
- E-mobility (15 new e-mobility solutions saving 95,5 T/year of CO_2)

Objective 2. *Smart Mobility*

- Multimodal mobility station/Central information pillar displaying all mobility options [*Munich, Vienna*]
- Smart street lamps with sensors and Wi-Fi access points [*Munich, Vienna*]
- E-car sharing fleet [*Munich, Vienna, Lyon*]
- Public bike rental system [*Munich*]
- E-Charging stations [*Munich, Vienna, Lyon*]
- E-trikes/E-cargo bikes sharing for transport of goods [*Munich, Vienna*]
- District sharing boxes- 24/7 pick-up stations (smart delivery/freight distribution/neighbourhood goods exchange) [*Munich, Vienna*]
- Electric delivery (E-vans & fast chargers) [*Vienna*]
- Munich Smart City app [*Munich*]
- E-bike and E-cargo sharing app [*Vienna*]
- Feasibility study of an E-taxi fleet station in Simmering [*Vienna*]
- Siemens E-fork lifts and charging stations for employees [*Vienna*]
- Autonomous driving E-shuttles (Navly) [*Lyon*]

Objective 3. *ICT*

- Smart Data Platform; Infrastructure:60 ‘intelligent’ lampposts for lighting and IoT (Living Lab) [*Vienna*]
- Complex monitoring system [*Munich*]
- Metropolis data platform [*Lyon*]
- Urban Living Lab SIMmobile (local communication and participation hub) [*Vienna*]
- Central data management system (Data Platform - Fiware) [*Vienna*]
- Community Management System (Lyon)
- Lyon Confluence Community Management System (CMS)

Smart Solutions (Devices/Systems) useful for official statistics

Objective 1. *Smart Energy*

- Electric-Renewable Energy Sources (energy statistics)

Objective 2. *Smart Mobility*

- Smart led street lighting with sensors
- E-bike sharing app
- E-cargo sharing app
- E-trikes/E-cargo bikes sharing for transport of goods
- District sharing boxes - 24/7 pick-up stations

Objective 3. *ICT*

- Lyon Confluence Community Management System (CMS)

Data output

- Deliverables provided at: https://www.smarter-together.eu/file-download/download/public/433
  Smart Data Platforms - raw data/KPI list not available on project website/public deliverables.
  - Smart Data Platform Munich D4.4.1
  - Lyon Confluence Community Management System (CMS) D3.4.4
  - Central data management system Vienna D5.4.2
- Compiled from the project's data set description:
  - Data of parking bookings and emissions
  - Traffic lights, traffic *(also relevant for smart mobility?)*
  - Raw data from electricity use for street lighting
  - Raw data from waste handling
- Sensor data from street environment *(also relevant for smart mobility?)*
- Shared electric mobility in Lyon-Confluence D3.5.1
  KPIs:
  - *number of electric vehicles charged at the stations*
  - *average number of kilometres travelled by a vehicle using charging stations*
  - *estimation of the contribution of charging stations to the reduction of CO2 emissions*
  - *distance travelled by each Navly shuttle*
  - *number of passengers carried by the shuttles*
  - *electric consumption of the Navly charging stations*

9.9 RemoUrban

All Smart Solutions

Objective 1. *Improve energy efficiency for sustainable districts and the built environment; Increase the thermal and electrical energy distribution and efficiency; Redirecting energy production towards Renewable Energy Sources;*

- In each of the retrofitted property a centralized intelligent control system will be installed, which will optimize energy use and storage to suit predicted demand profiles.
- Charging infrastructure [*Valladolid, Nottingham, Tepebasi/Eskisehir*]

Objective 2. *Transport sustainability; Develop innovative technologies;*

- Multimodality passenger mobility based on RFID or smart cards [*Valladolid, Nottingham, Tepebasi/Eskisehir*]
- Intelligent transport systems (app, traffic models, smart devices on vehicles)/ Road Management Systems [*Valladolid, Nottingham, Tepebasi/Eskisehir*]
- Traffic flow control (smart traffic lights system, smart speed limiter system, dynamic route planner, dynamic parking place allocation) [*Miskolc*]
- Public car sharing [*Valladolid*]
- Smart traffic lights

Objective 3. *Reduce GHG emissions in urban areas; Reduce CO2 emissions to 50%*

- Low carbon solutions for thermal energy supply and optimized electric facilities by means of decentralized electricity generation and smart grid management will be deployed in order to achieve near zero energy and zero emission districts
- Last mile deliveries E-vehicles [*Valladolid, Nottingham*]
- E-vehicles (Full/Hybrid) [*Valladolid, Nottingham, Tepebasi/Eskisehir*]

Objective 4. *Connect key information sources with city monitoring systems (sensors, people); with city 'life-lines' infrastructures (transport, power, water, and communication) to build city resilience against incidents and crisis*

- City Information Platform [*Valladolid, Nottingham, Tepebasi/Eskisehir*]
- Mobile ITS Technology [*Valladolid, Nottingham, Tepebasi/Eskisehir*]
- Inter-operability and data protocols between city domains [*Valladolid, Nottingham, Tepebasi/Eskisehir*]

Smart Solutions (Devices/Systems) useful for official statistics

Objective 1. *Improve energy efficiency for sustainable districts and the built environment; Increase the thermal and electrical energy distribution and efficiency; Redirecting energy production towards Renewable Energy Sources;*

- Monitoring tools for energy [*energy statistics; electric energy consumption statistics; energy consumption optimization models*]

Objective 2. *Transport sustainability; Develop innovative technologies;*

- RFID or smart cards [purpose: *mobility statistics; traffic optimization models*]
- Smart traffic lights [purpose: *traffic statistics; mobility statistics; road accidents statistics*]
- Mobility app [purpose: *mobility statistics*]

– Smart devices on vehicles [purpose: *traffic statistics; mobility statistics*]

Objective 3. *Reduce GHG emissions in urban areas; Reduce CO_2 emissions to 50%*

– Monitoring tools for energy [*energy statistics; electric energy consumption statistics; energy consumption optimization models; energy balance*]
– Smart devices on EV measuring electricity consumption [purpose: *electric energy consumption statistics; emissions accounts/models*]

Objective 4. *Connect key information sources with city monitoring systems (sensors, people) with city 'life-lines' infrastructures (transport, power, water, and communication) to build city resilience against incidents and crisis*

- City Information Platform [purpose: *traffic statistics; mobility statistics; road accidents statistics, light utility vehicles (LUV) freight transport statistics; energy consumption models; policy intervention evaluation models*]

Data Output

City Information Platform - raw data/KPI list not available on project website/public deliverables.

9.10 Ruggedised

All Smart Solutions

Objective 1. *Energy reduction/Smart Thermal Grid*

- Thermal Energy Storage
  - Geothermal heat/cold storage and heat pumps [*Rotterdam*]
  - Geothermal heat/cold storage [*Umeå*]
- Extracting/ Collecting Thermal Energy
  - Thermal energy from waste streams [*Rotterdam*]
  - Surface water heat/cold collection [*Rotterdam*]
  - Pavement heat/cold collector [*Rotterdam*]
- Thermal Exchange And Connected Buildings
  - Peak load variation management and peak power control [*Umeå*]
  - Heat exchange-connection of buildings to district heating network [*Glasgow*]
- 100% Renewable Energy Business Model
  - Smart City connection to 100% renewable energy [*Umeå*]

Objective 2. *Reducing environmental impacts developing e-mobility solutions / Smart Electricity Grid*

- E-Charging Infrastructures
  - Smart charging parking lots [*Rotterdam*]
  - EV-charging infrastructure hub [*Umeå*]
  - Surplus power storage in e-vehicle charging hub [*Glasgow*]
  - Innovative connection to renewables and storage [*Glasgow*]
  - Intelligent LED street lights with integrated EV charging functionality [*Glasgow*]
- Electric Buses
  - Optimising the E-bus fleet [*Rotterdam*]
  - Energy optimised electric Bus Rapid Transit-station (BRT) [*Umeå*]
- Renewable Energy Storage
  - EV-charging hub battery storage in car parks [*Glasgow*]
  - Renewable Energy Sources (RES) generation and storage for mobility [*Rotterdam*]
  - Optimisation of integration of near-site Renewable Energy Sources (RES) [*Glasgow*]
- Green Parking
  - Energy-efficient land use through flexible green parking [*Umeå*]

Objective 3. *Energy Management And ICT*

- Energy Data Management
  - Energy management in buildings [*Rotterdam*]
  - 3D city operations model [*Rotterdam*]
  - Long-range wireless network [*Rotterdam*]
  - High performance servers to heat homes [*Rotterdam*]
  - Smart waste Management [*Rotterdam*]
  - Intelligent building control and end user involvement [*Umeå*]
- Smart Open Data Platform
  - Smart open-data city platform [*Umeå*]

        - Smart open-data decision platform [*Glasgow*]
    - Intelligent street lighting
        - Efficient and intelligent street lighting [*Rotterdam*]
        - Intelligent LED street lights with integrated EV charging functionality [*Glasgow*]
    - Demand-Side Management
        - Demand-side management technology in a university campus [*Umeå*]
        - Implementation of demand-side management technology in street lighting, domestic and non-domestic properties [*Glasgow*]

Smart Solutions (Devices/Systems) useful for Official Statistics

Objective 1. 100% Renewable Energy Business Model

- Smart City connection to 100% renewable energy [*purpose: energy statistics; energy balance; energy consumption models*]

Objective 2. Reducing environmental impacts developing e-mobility solutions / Smart Electricity Grid

- E-Charging Infrastructures
    - EV-charging infrastructure hub [*purpose: mobility statistics; energy statistics]*
- Renewable Energy Storage
    - Renewable Energy Sources (RES) generation and storage for mobility [*purpose: energy statistics*]

Objective 3. Energy Management and ICT

- Smart Open Data Platform
    - Smart open-data city platform [*purpose: energy statistics; energy consumption models; travel patterns*]
    - Smart open-data decision platform [*purpose: policy intervention evaluation models*]
- Demand-Side Management
    - Implementation of demand-side management technology in street lighting, domestic and non-domestic properties [*purpose: energy statistics; energy balance*]

Data Output

- <u>Smart Open-Data (city/decision) Platform: raw data/KPI list not available, yet.</u>

9.11 MySmartLife

All Smart Solutions

Objective 1. *Reducing CO_2 emissions and increasing the use of renewable energy sources. New services concerning electricity*

- Retrofitting Actions *[Nantes, Hamburg, Helsinki]*
- Renewable Energy Production and Management: photovoltaic (PV) projects *[Nantes*]
- Smart metering and smart controls *[Hamburg]*
- Solar power plants *[Helsinki]*
- Smart Street Lighting (the lampposts will be replaced towards smart street lighting, providing bicycle counters, LED lights and Wi-Fi.) *[Hamburg, Helsinki]*
- Energy Projects (Smart Heating Islands): 'Heating Island' includes a combined heat and power production (CHP) unit that serves heating power for a hotel, a new constructed cultural centre and 50 housing units. *[Hamburg]*
- High-Performance Residential Buildings: Integration of renewable energy systems and waste heat in the buildings as well as demand response. *[Helsinki]*
- Remote lighting management system *[Nantes]*
- E-bus fleet [Nantes, Hamburg, Helsinki]
- Automated electric minibus [Helsinki]
- E-car/E-bike sharing [Nantes, Hamburg]
- E-charging stations/points [Nantes, Hamburg]

Objective 2. *Involving citizens in the development of an integrated urban transformation strategy*

- Mobility hub (charging stations + shared mobility) [*Hamburg*]
- Innovative urban logistics [Nantes]
- Intermodal observatory ('an intelligent data-based tool allowing to track mobility data, monitoring the actions of the Mobility Plan define key mobility indicators and provide data to Nantes open data portal and urban platform') [Nantes]

Objective 3. *Increasing the digitalization of the cities through the urban platforms*

- Started in 2017: Energy DataLab ('Enedis platform'): displays electricity consumption data for a sample of 300 buildings and public lighting cabinets. Dynamic electricity data platform hosting data sets collected from the smart electric meters (amongst other data collected). As an open platform, it will provide those data to the urban platform through open Application Programming Interfaces (APIs). [Nantes]
- Urban Platform will be supplemented with the DTAG (T-Systems, AG) Smart City Ecosystem to form a 'System of Systems' following the DIN Spec 91357 on 'Open Urban Platforms'. It's a new kind of web service for real-time sensor information. Following Smart City applications will be connected to the smart middleware of the Deutsche Telekom's Smart City Ecosystem: smart street lighting, intermodal routing and smart grid [Hamburg]
- Urban Platform: new open data generated in mySMARTLife and sourcing the heat leakage images of building facades. 3D city model and IoT platforms are implemented and data available from the sectorial ICT systems will be utilised. [Helsinki]

- to develop and integrate urban platform extensions into Metropole's existing information system in order to optimise public policies and provide more efficient public services or to provide citizens with new services [Nantes]
- Planned: combined 'platforms': and other Smart City Platforms, e.g. to the partner Lighthouse Cities [Hamburg]

Smart Solutions (Devices/Systems) useful for Official Statistics

Objective 1. *Reducing CO_2 emissions and increasing the use of renewable energy sources. New services concerning electricity*

- Smart street lighting with sensors [*purpose: traffic statistics; mobility statistics; road accidents statistics; electric energy consumption statistics; energy consumption optimization models*]
- Energy DataLab, uses data from about 150.000 smart electric meters [*energy statistics; energy balance; emissions accounts/models*]

Objective 2. *Involving citizens in the development of an integrated urban transformation strategy*

- Intermodal observatory [purpose: *traffic statistics; mobility statistics; road accidents statistics, light utility vehicles (LUV) freight transport statistics*]
- Mobility hub [purpose: *mobility statistics; emissions statistics]*

Objective 3. *Increasing the digitalization of the cities through the urban platforms*

- Urban Smart Platforms [purpose: *traffic statistics; mobility statistics; road accidents statistics, light utility vehicles (LUV) freight transport statistics; energy consumption models; policy intervention evaluation models*]

Data Output

Urban Smart Platforms

Link:https://mysmartlife.eu/fileadmin/user_upload/Deliverables/D2.16_Open_Specifications_framework.pdf

9.12 Sharm-LLm/Sharing cities

All Smart Solutions

Objective 1. *Pilot energy efficient districts / Shift thinking irreversibly to local renewable energy sources*

- Building Retrofit (improved heating systems, energy efficient lighting, Installing renewable energy sources such as solar panels to provide locally generated energy with less emissions and greater energy security, Installing devices which can track energy use, temperature and humidity in the home) [*Lisbon, London, Milan*]
- Sustainable Energy Management System (SEMS) [*London, Lisbon, Milan*]
- Smart Lamp Posts [*London, Lisbon, Milan*]
- District Heat Network: a heat network to supply social housing units with low cost and low carbon energy *[London]*

Objective 2. *Make at least 10% of local citizens chose electric over fossil fuel vehicles / Reduce CO_2emissions*

- Smart street lighting with sensors [*London, Lisbon, Milan*]
- E-charging stations
- E-car/E-bike sharing
- Smart parking
- E-logistic
- Mobil.me platform [*Lisbon*]

Objective 3. *Providing an interactive representation of a sustainable model with the contribution of citizens and the data acquired by IoT*

- Urban Sharing Platform (USP): the main purpose of the USP is to aggregate data and control from a wide variety of devices and sensors, store and process the data, and support visualization of the information to the city and citizens, which enables better use of the city resources. [*London, Lisbon, Milan]*
- Smart Lamp Posts: to provide lighting to the city more efficiently. Integration with other smart city measures through sensors. People and things can communicate with data concentrators, installed on public lighting streetlights, so they can always be connected. The lamp post will not only be devoted to illuminating the street or sidewalk where it is installed but will become a key node for network infrastructure that transforms and improves community services. Monitoring of environmental services, security, smart mobility, parking and much more can be managed with applications from both citizens and municipalities. [*Lisbon, London, Milan*]

Smart Solutions (Devices/Systems) useful for Official Statistics

Objective 1. Pilot energy efficient districts / Shift thinking irreversibly to local renewable energy sources

- Sustainable Energy Management System (SEMS) [*purpose: energy statistics; energy consumption models*]

Objective 2. Make at least 10% of local citizens chose electric over fossil fuel vehicles / Reduce CO_2 emissions

- Smart street lighting with sensors: [purpose: *traffic statistics; mobility statistics; road accidents statistics*]

Objective 3. Providing an interactive representation of a sustainable model with the contribution of citizens and the data acquired by IoT

- Urban Sharing Platform (USP) [*purpose: energy statistics; mobility statistics; policy intervention evaluation models*]
- Smart Lamp Posts [*purpose: mobility statistics; traffic statistics; road accident statistics]*

Data Output

Urban Sharing Platform (USP)
(http://nws.eurocities.eu/MediaShell/media/Urban_sharing_platform_requirements.pdf):
it is a logical collection of technical components, capabilities and processes which provides functions and services that enable a smart city. Its purpose is to aggregate data and control functions from a wide variety of devices and sensors, store, process, correlate the data and present information to the city and citizens which enables better use of the city resources and may provide support for innovative service verticals.

Case Study 3

1 Introduction

1.1 Goals

The many sensors installed in so called "smart cities" allow cities to collect regularly updated data about air pollution, traffic and noise. Yet these data alone are of limited interest to guide local decision-makers in the elaboration of their public policies. On the contrary, cross referencing city data with socio-economic data collected by NSIs allows shedding light on new and accurate issues.

Exposure of populations to air pollution is a major public health issue. Its involvement is proven in cases of diminished respiratory capacity, asthma and inflammation of the lungs, increased vulnerability to respiratory infections and premature death (Kihal et al., 2015).

The European Environment Agency published a study at the end of 2018 entitled "Unequal exposure and unequal impacts: social vulnerability to air pollution, noise and extreme temperatures in Europe". This study highlights a global trend towards overexposure of disadvantaged populations, which are also the most sensitive to the effects of pollution. The study also highlights the lack of information at a fine spatial scale. Cross-referencing air pollution data with NSI's geo-localized data on socio-economic characteristics of the inhabitants makes it possible to identify the areas that accumulate high exposure to air pollution, and underprivileged inhabitants. This allows decision-makers to precisely target areas where priority action is needed.

1.2 Workflow and first results

During the first 10 months of workpackage L, ISTAT and INSEE have followed the subsequent steps : (i) finding a pilot city, both in France and in Italy, equipped with pollution sensors, for which we had access to data; (ii) identifying the precise partner to involve in the project; (iii) performing an academic literature review, thereby defining the methodology of the study (iv) analyzing the data gathered by the sensors; (v) analyzing the socio-economic data available at the NSIs and identifying the variables of population and dwellings most useful for the study;

The Nice metropolis (south of France) hosts the IMREDD (Mediterranean Institute for Risks, Environment and Sustainable Development). This centre is co-founded by the metropolis of Nice and the University of Nice Côte d'Azur. It compiles all the data collected in the city of Nice by the many sensors for traffic, pollution, noise, etc. Thanks to a dedicated platform (the Smart City Innovation Center), the institute makes the data available to researchers or companies with a project related to the Nice area.

The alignment of the timeline between INSEE's explorations of new data sources, and IMREDD's investigations for finding new partnership in order to valorize their data has enabled INSEE's unit for territorial analysis and IMREDD's department for innovation and strategic partnerships (leaded by Paulo Moura) to work concretely on the subject and to carry out a first study on the links between socio-economic vulnerability and exposure to air pollution in the city of Nice. This study tends to show that in the Nice city, people of higher socio-economic status are significantly less exposed to air pollution than people of medium and low socio-economic status.

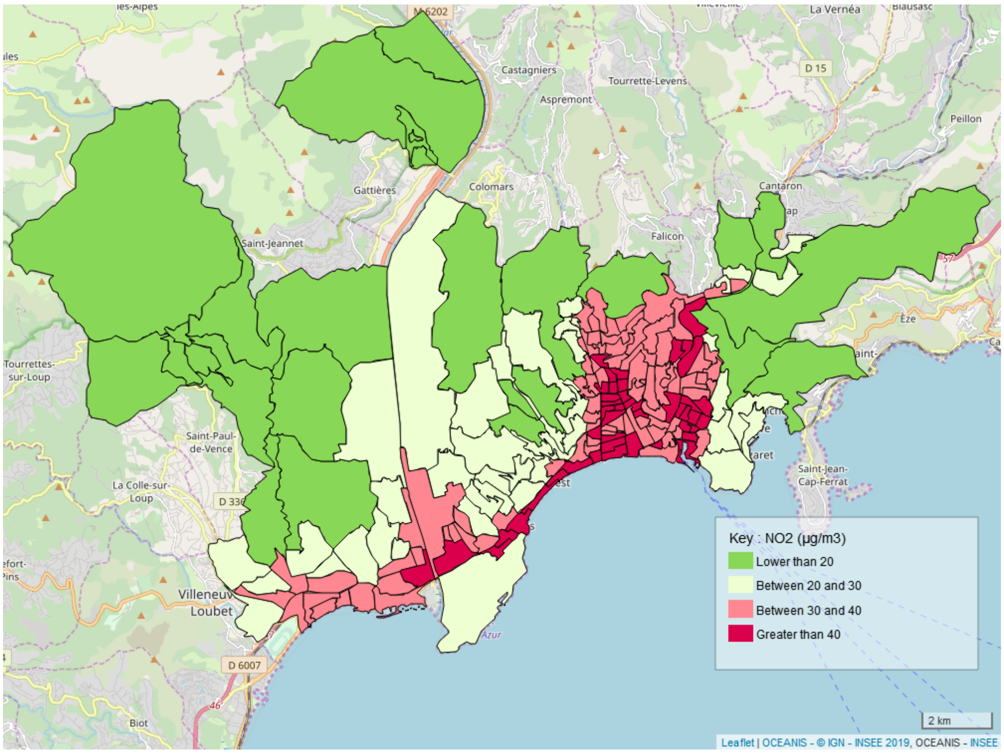

Figure 8: NO2 mean annual concentration in the Nice metropolis

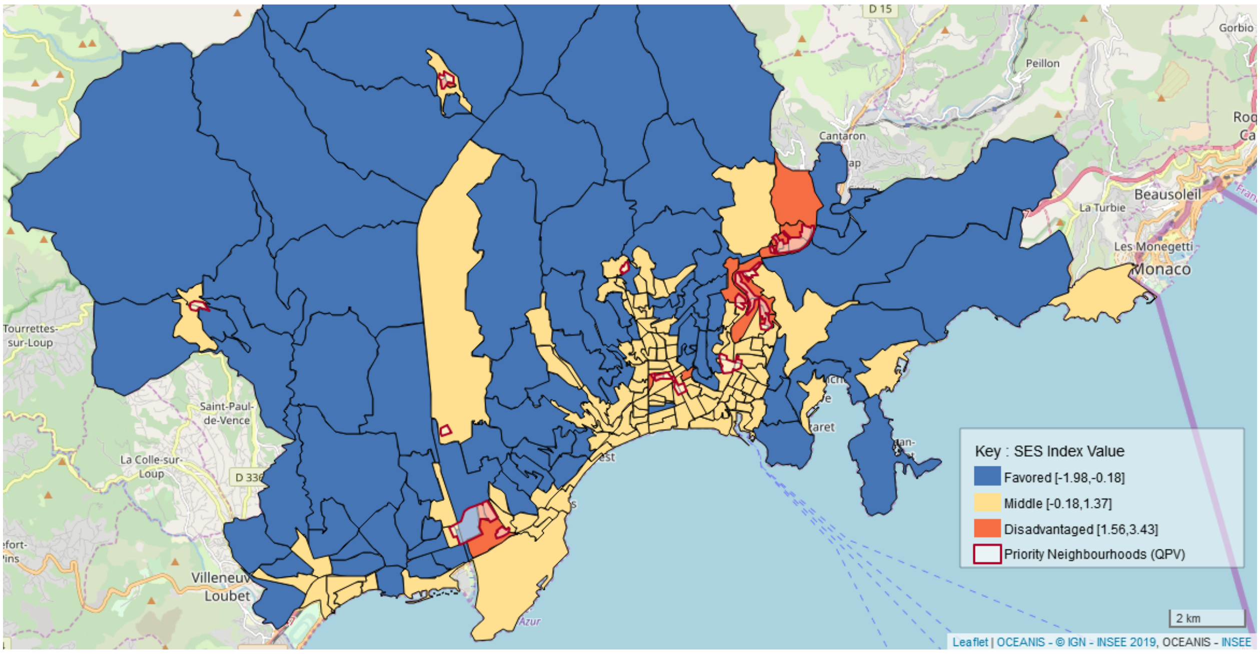

Figure 9: Socio-Economic Index Value in the Nice metropolis

1.3 Following steps

INSEE

For this first step of the study, IMREDD has mainly brought its knowledge of the particularities of the city (socio-economic context, ongoing public policies about mobility and roads with high daily traffic). The pollution data that were used in this study are indeed available online, as open data. Now that the areas that cumulate underprivileged inhabitants and high pollution have been identified, IMREDD will enrich the analysis with an understanding of the reasons for a high pollution in these areas.

There the value added for an NSI to work with a Smart City will be the highest. Indeed, IMREDD will produce information on car traffic in the most underprivileged areas. It will also add in its mobility modelization software, INSEE's data about the working place of Nice's inhabitants, in order to make some commuting simulations. Lastly, since IMREDD works jointly with the Nice Metropolis, which is responsible for the installation of pollution sensors, the Nice Metropolis will use our study to identify areas where a pollution sensor should be placed in order to make a more precise analysis. This is therefore a very promising collaboration between people which make the analysis (statisticians from INSEE and IMREDD), and the city which is in charge of data collection and which is also the public policy maker.

In the mid-term, the goal is to show other cities the results of our study. We hope to convince cities which were not eager to share their sensor data with an NSI that it may be in their own interest to do so. We could then reproduce our analysis for other cities and compare the results, which would be even more helpful to local decision maker trying to address social inequalities in health linked to air pollution.

Another axis of improvement for the mi-term is to study the socio-economic characteristics of people exposed to pollution during day time. For the moment, we restrict indeed our analysis to the place where people are domiciled, yet they can be exposed to other levels of pollutions depending on where they spend the day. Moreover, air pollution values are subject to large fluctuations during the day. This would be an interesting application for the work conducted by other groups at INSEE, about the estimation of day-time population.

ISTAT

ISTAT has participated in this case study with the aim of learning the INSEE experience and applying it to its context. In this sense, the next steps certainly will be to exploit the experience gained by INSEE colleagues, and to continue in defining the partnership with Regional Agency for Environmental Protection. This last step represents the essential point for this study to have a possible follow-up in the next European projects. The partnership will allow Istat to have a direct access to the micro-data that are neither published nor accessible on Agency's web site. Moreover, the partner has a fundamental dominion knowledge to carry on the case study.

1.4 Recommendations

Many smart cities outsource data collection to private companies that deliver them as a black box, which contains only aggregated data. These aggregated data are of limited interest for NSIs. One crucial point for launching a fruitful collaboration with a smart city, is therefore to **find a pilot city**, which would allow NSIs to have **access to raw data**. This access can be direct, or also indirect if the partner city does the analysis on the raw data and send to the NSI useful aggregated data. The main point is that there has to be some engineers in the pilot city or researchers who have access to the sensors data, and work hand in hand with the NSI for exploiting these data.

The second recommendation would be to define a project with a **simple and well documented issue, linked with a topical public policy**. This will allow local decision makers to be directly involved in the results of the study, and therefore more eager to ease the process of access to the data. A well-defined topic and a relatively simple methodology will allow the resulting study to be published within a reasonable time. Therefore, the proof of concept of the interest in combining NSI socio-economic data of NISs with data collected by the smart city will quickly be proven.

2 Study background

Exposure of populations to air pollution is a major public health issue. Its contribution is proven in cases of decreased respiratory capacity, asthma and inflammation of the lungs, increased vulnerability to respiratory infections and premature death (Kihal et al., 2015). One of the components of air pollution, fine suspended particulate matter (PM2.5), has been recognized as causing nearly 391,000 premature deaths in the 28 EU member states, and up to 422,000 in the 41 European countries, for the year 2015 alone (EEA, 2018). Together with Particulate Matter PM10, they are among the pollutants that have received the most attention from the European office of the World Health Organization (WHO), in addition to nitrogen dioxide (NO_2), which is known to be a good indicator for traffic-based air pollution (Padilla et al., 2014).

Beyond the harmful effects of environmental pollution on health, the literature highlights the uneven distribution of its impacts on individuals. Social inequalities in health (ISS) are defined as "any relationship between health and belonging to a social category" (Inpes, 2012).

Many publications show that children, the elderly and disadvantaged populations are more sensitive to pollution. Indeed, their lifestyle often combines poor nutrition, inadequate access to care and a greater predisposition to certain factors, such as stress (Khreis et al., 2017).

The other side of the social inequalities in health linked to pollution - much less documented - is the greater exposure of disadvantaged populations. The European Environment Agency published a report in 2018 on the differences in exposure to various environmental risks (air pollution, noise pollution and extreme temperatures) of European populations, according to their socio-economic category. It shows that at the European level, the poorest regions in the eastern and south-eastern parts of the continent tend to be the most polluted (in terms of fine PM particles). On the contrary, higher NO_2 levels have been found in the richest regions due to high traffic and industrial activity, but within these areas, the populations exposed to NO_2 are the poorest. This is also found at a more local level, and more particularly in urban areas where the most disadvantaged populations appear to be the most overexposed (EEA, 2018). However, other studies conducted in Europe on a finer geographical scale, do not always converge towards the same conclusions (Padilla et al., 2014).

In France, only a few studies on this subject are carried out by a team of researchers from the Ecole des Hautes Etudes en Santé Publique (School for Advanced Studies in Public Health), whose objective is to "explore the contribution of a certain number of environmental exposures to social inequalities in health" through the "Equit'Area" project. These studies have focused, in particular, on the major cities of Paris, Lyon, Marseille and Lille. The main results show differences in the correlations between exposure to pollution and the socio-economic characteristics of individuals. The more prosperous districts (with a high proportion of executives and high average incomes) are more exposed to air pollution (NO_2) in Paris, while in Lille and Marseille it is the disadvantaged population (with a high proportion of immigrants) that is affected the most. In Lyon, it is the middle class that are most exposed to NO_2 (Padilla et al., 2014). Another study in Strasbourg highlighted the existence of a link between social disadvantage and exposure to air pollution (the considered pollutants are NO_2, PM10 and carbon monoxid (CO)) (Deguen, 2013).

In the city of Nice, which is the subject of our pilot project study, no such studies have been carried out. However, a publication in the scientific journal "Environmental Health" (Padilla et al., 2016) is methodologically similar and also addresses the issue of exposure to environmental pollution. This study links three factors: socio-economic characteristics, access to care and exposure (positive or negative) to environmental factors, with the infant and neonatal mortality rate at the IRIS level (inframunicipal statistical breakdown produced by INSEE, with at least 2000 inhabitants). It reveals that some areas, which are more socially disadvantaged, are more exposed to environmental pollution and present a higher risk of infant and neonatal mortality, but still have access to proper care combined with good transport services. As a result, the populations residing in these neighbourhoods are not considered to suffer from social inequalities in health linked to pollution.

Public authorities are increasingly implementing public policies to combat air pollution and thus preserve the health of their citizens. However, some interventions may unintentionally contribute to increased social inequalities in health (Kihal et al., 2015).

This risk is particularly high in cities, where new means of public transport are introduced. The commissioning of a new tramway will have a positive impact on the quality of life due to the reduction in car traffic and easier access to places of employment. However, this will lead to an increasing valuation of real estate in the neighbourhoods, which in turn will see the most disadvantaged populations leave for the benefit of the wealthiest (Faburel, 2008). There are thus strong correlations between policies to protect the environment and improve mobility and the evolution of property prices. The fact that the most disadvantaged, financially constrained populations favour the low price over the neighbourhood's quality of life, highlights the influence of the real estate market on social and environmental inequalities. For example, in the agglomerations of Lille and Marseille, prices in favoured (and less polluted) areas are about 40% higher on average than those in disadvantaged (and more polluted) areas (Padilla et al., 2014).

The European Environment Agency, in its latest report (EEA, 2018), makes the same alarming observation that the policies pursued have rather benefited the privileged populations. The agency looked in more detail at the lack of consideration for the most socially deprived, and in particular highlighted the lack of actions targeted at particular groups of people, as well as the absence of a fund dedicated to socio-environmental inequalities.

In 2007, the metropolis of Nice opened a first tramway line, which was extended in 2013, and a second line in 2019. The objective is to boost access to jobs and improve the quality of life in the city. One of the expected consequences of the opening of this line is the reduction of air pollution in the centre of Nice. At the same time, improving access to jobs and improving the quality of life in central neighbourhoods should lead to higher housing prices and therefore possibly to a shift of disadvantaged populations to other potentially more polluted neighbourhoods. The objective of the planned study is to describe the evolution of the socio-economic profiles of individuals exposed to high air pollution over a period corresponding to the commissioning of tramway lines in Nice. The metropolis of Nice will be kept regularly informed of the progress of this study, which will shed new light on the public policies it is undertaking. In addition, the city has several mobile pollution sensors, which can be placed in the critical areas identified by the study.

3 Compared scoping data

	INSEE	**ISTAT**
Pilot City	Nice (FR)	Rome (IT)
Population	342 637 inh. (5th in France)	2 873 000 inh. (1st in Italy)
Surface	72 km2	1 285 Km2
Mean household income	15 563 €	37 547 €
Project's stakeholders	● IMREDD (depends from Nice University and the Nice metropolis)	● ARPA Lazio (Regional Agency for Environmental Protection)
Number of pollution sensors	3 fixed + 5 mobile	13 fixed
Pollutant studied	● PM10 ● PM2.5 ● NO2	● PM10 (Particulate Matter) ● NO2 (Nitrogen Biosside) ● O3 (Ozone) ● CO (Carbon Oxide) ● SO2 (Sulfur Oxides)
Temporal frequency	1 hour	1 hour
Temporal period	5 year	1 year
NSI's source	● census data ● fiscal data	● census data
Geographical scale available	35 m2	1 Km2
Geographical scale of the study	1km2	

Table 1 : comparison of scoping data of the italian and french projects

4 The INSEE's project in detail

4.1 Partnership

IMREDD, attached to the University of Nice Côte D'Azur, has created the Smart City Innovation Center. This Center - unique in France in terms of its operation and the collected data - centralises data from sensors installed by the Nice metropolis and other private or public partners. The researchers at the Center format the data to facilitate further analysis. IMREDD then enters into partnerships to promote this data in projects that are useful to the territory of the Nice metropolis.

INSEE and IMREDD have written a master agreement, stating that "the objective of the partnership between INSEE and IMREDD is to improve knowledge of the territory of the metropolis of Nice, and more generally of the Maritime Alps, in order to inform local decision-makers. The partners will mutually enrich each other by sharing knowledge on their respective databases, by discussing the choice of the most relevant issues to be studied, as well as by developing methodologies to link their data".

This master agreement is currently being proofread by the legal departments. It is accompanied by an application agreement, which details the issue of the first study resulting of this partnership: "links between socio-economic status and exposure to air pollution in the Nice area". The study will be published as a four page article, in INSEE's editorial line. Although the type of collaboration is entirely new, it was formalized in an agreement which is based on the traditional conventions of partnership study with a local public actor. The main difference is the timetable (much longer in our case : one and a half year dedicated to the study, versus six months for a traditional study) and the fact that the partner will bring not only its knowledge of the territory and the public policy, but also detailed data collected by the sensors which he is in charge of installing.

IMREDD has also links with AtmoSud, which is the regional agency for air quality monitoring. AtmoSud publishes in open data annual information about pollutants concentrations. The detailed data are available on demand. Experts from AtmoSud will help us to interpret our results.

4.2 Data

Socio-economic data

One added value of our study in comparison to the existing literature is due to the richness of INSEE's geolocalized data about dwellings and their inhabitants. This information is extracted from a source called "Fideli". The Demographic File on Housing and Individuals (Fidéli) is the result of a project launched by INSEE in 2011 to create a statistical database on housing and individuals. In recent years, administrative data have been increasingly used in the production of public statistics, which has encouraged this project to develop.

The Fideli database is obtained from tax files (mainly those of the housing tax), combined with additional data related to the cadastral parcels on which the dwellings are built. This database contains information on dwellings: geographical coordinates, nature of the owner (individual, company, HLM), tax status of occupation (owner, tenant), main or secondary residence, number of floors, years of construction as well as different physical characteristics: surface area, number of living rooms, bedrooms, kitchens, bathrooms, presence of an elevator, a garage or a box. Detailed information on the inhabitants of the dwellings is also available: sex, years and places of birth, marital tax status, links with the tax registrant and information on income (wages, unemployment income, retirement pensions, etc.).Finally, a variable indicates whether individuals moved between year N and year N-1.This database was used for the year 2016.

The other source used by INSEE to characterize people's socio-economic vulnerability is the national census. In municipalities with more than 10,000 inhabitants, a sample of 8% of the population is selected every year, and the results are calculated after five years on the basis of a sample composed of 40% of the population. The results are divided into a main exploitation, covering all the bulletins collected but only some of the variables, as well as a complementary exploitation which is carried out only on some of the respondents and provides details on occupations and socio-professional categories, sectors of economic activity, employment and the family structure of households.
In order to guarantee the statistical quality of the results, we conducted our study on a geographical area which fits the census' sample plan: IRIS (inframunicipal statistical breakdown, with at least 2000 inhabitants).
The most relevant variables were chosen thanks to a statistical method, detailed in part 2.5.

Pollution data
The pollution data come from the association for monitoring air quality in the Provence-Alpes-Côte d'Azur region, called Atmosud. They are available on the DataSud website, which is a regional open data platform. These are annual models, in micrograms per cubic metre (µg/ m^3) at the scale of the region, carried out on the basis of information collected by stations (about sixty, most of them recording one to three pollutants and this can go up to ten) and taking into account a multitude of external elements (weather, wind, proximity of plants...). Modelling is carried out for the main air pollutants, which are NO_2, fine particulate matter (PM10 and PM2.5) and ozone (O3). The accuracy corresponds to squares of 25 m side, except for ozone where it is only related to 1 km², which makes it difficult to use the results for this study. This data is available in Geotiff format, allowing geographic information (projection, coordinate system) to be added to an image. There is thus one ".tiff" file per pollutant for each year from 2013 to 2017.

4.3 Methodology

We relied on the methods used by the Equit'Area team to build a social disadvantage index. The economic and social status of a group of individuals is determined on the basis of socio-economic variables that cover the following fields: income, employment, education, household composition, immigration status, previous mobility, housing characteristics. This method makes it possible to go beyond the simple income variable.
From the census and the fiscal data sources, initial variables are selected for their explanatory potential in terms of people's economic and social living conditions. Then, an aggregated value per IRIS, the statistical unit used for the rest of the study, is calculated for each of them (IRIS is an inframunicipal statistical breakdown, with at least 2000 inhabitants).

Here is the list of the 53 variables that have been determined at the IRIS scale:

Area	Variable (in percent)	Source
Family and household	individuals under 25 years of age in the total population individuals over 65 years of age in the total population individuals living outside the household individuals living in single-parent families in the total population people living alone in the total population	Census 2015 - individuals
Immigration and mobility	immigrants in the total population* foreign individuals in the total population relocation since the previous year	Census 2015 - individuals Fidéli 2015 – individuals

Education	without diplomas (or very low level) in the population aged 15 or over vocational qualification graduates in the population aged 15 or over bachelor's graduates in the population aged 15 or over higher education graduates in the population aged 15 years or over children aged 6-15 years in the population students in the population aged 15 and over	Census 2015 - individuals
Employment	farmers in the labour force managers in the labour force workers in the labour force self-employed entrepreneurs in the labour force individuals with stable employment in the labour force individuals with precarious employment in the labour force workers in the total population working men in the total male population working women in the total female population unemployed in the labour force individuals who have been looking for work for more than a year in the labour force foreign unemployed in the labour force unemployed 15-24 years old in the labour force unemployed people aged 50 and over in the labour force unemployed in the male labour force unemployed in the female labour force	Census 2015 - individuals
Income	median salary of individuals who receive a salary households considered poor (at 60%) average amount of housing subsidies paid per household average amount of minimum social subsidies per household average amount of family benefits paid per household median standard of living per household	Fidéli 2015 - individuals Fidéli 2015 - income
Housing	principal residences built before 1970 principal residences built after 1990 principal residences less than 40m². principal residences superior to 120m². principal residences with a parking / garage social housing (HLM) average number of individuals per room households with more than one person per room households who do not own the housing individual housing multiple housing car-free households households with 2 or more cars average length of residence in the same dwelling in Iris	Census 2015 - housing
	principal residences without bath or shower principal residences without toilets principal residences without central heating	Fidéli 2016 - housing

Table 2 : first selection of variables with an explanatory potential in terms of people's economic and social living conditions

Redundant group: active population b

Redundant group: unemployed population c

Redundant group: social assistance / poverty d

Redundant group: foreign population e

Redundant group: financial resources f

*Variables selected from each group of redundant variables **

Four steps were successively carried out, with the final objective of preserving only the variables that provide the best definition of the socio-economic context of individuals.

Step 1: Study of redundant variables

Among the variables detailed in table 1, many express the same concept and are strongly correlated. The correlated groups are identified by letters b to f in table 1.

For each of these groups, a Principal Component Analysis (PCA) is performed, and the variable most correlated to the first component is retained because it corresponds to the best representation of the group of variables.

For example, for the "active population" group, the variable defining the share of workers in the total population, contributes most to the main axis, and must therefore be retained at the expense of the others. The variables selected for the other groups of redundant variables are:

- the share of unemployed in the labour force
- the share of foreign individuals in the total population
- the average amount of housing benefits paid per household
- the median standard of living per household

Step 2 : Variable selection

Again, a PCA is performed on all 40 remaining variables at the end of the first step. The objective is to select variables for which the contribution to the first axis is greater than the average contribution. Indeed, the first component of the PCA on socio-economic variables can be interpreted as a component of "economic and social status". 18 variables are selected at this step.

Step 3 : Index construction

The last PCA on the 18 retained variables makes it possible to distinguish between variables contributing to a better socio-economic level and those that are more likely to be sign of disadvantages.

Step 4 : Creation of socio-economically homogeneous classes

- Once the value of the disadvantage index is obtained for each of the IRIS, a hierarchical bottom-up classification is performed to group socio-economically similar IRIS into classes.

18 variables were selected for the index construction:

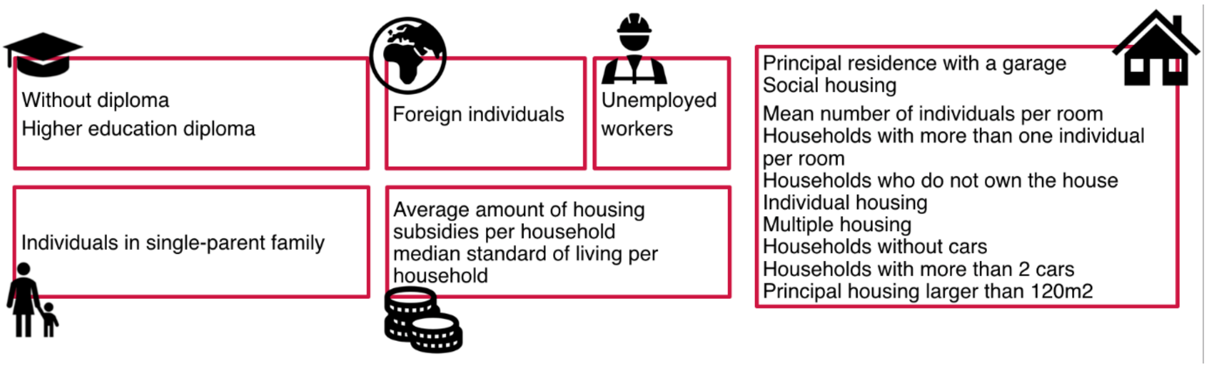

Figure 10 : variables selected for the construction of the socio-economic index

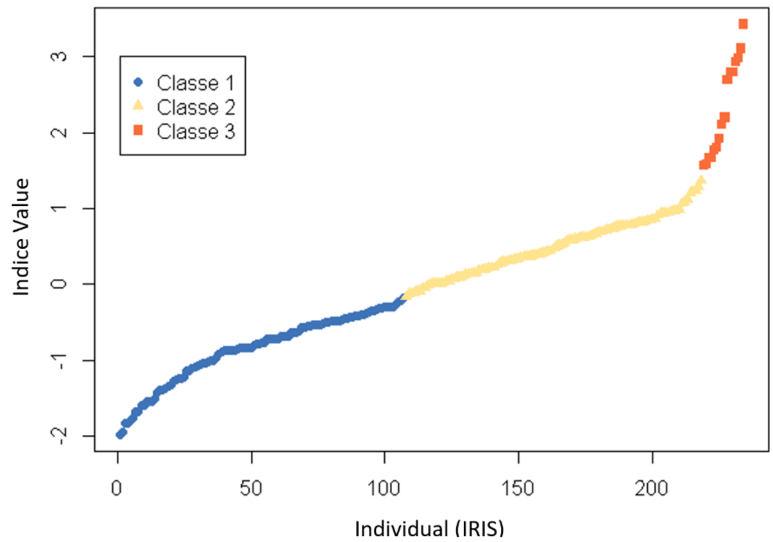

Figure 11 : Index value per IRIS

The graph in figure 3 presents the value of each individual's socio-economic index (in the statistical sense, it is indeed the IRIS) sorted in ascending order. It corresponds to the individual's coordinate on the first component of the last PCA performed, which was centered and reduced. The higher the index, the more disadvantaged the corresponding IRIS is considered to be. Thus, we see that most IRIS belong either to the most favoured class (in blue) with a population of 107, or to the intermediate group (in beige, 111 IRIS). For the first group, the values are between -1.97 and -0.18, while those of the second class are up to 1.36. The last group, the most disadvantaged IRIS (in orange), has only 16 individuals. The index values range from 1.56 to 3.43 for the most disadvantaged. It is in the intermediate class that individuals are the least dispersed, with a standard deviation of 0.37 for a median of 0.47. The standard deviation is 0.45 in the wealthiest class, with a median value of 0.78, and finally the most disadvantaged class has a higher dispersion since its standard deviation is 0.63 and its median index is 2.1.

<u>Construction of the pollution data base</u>

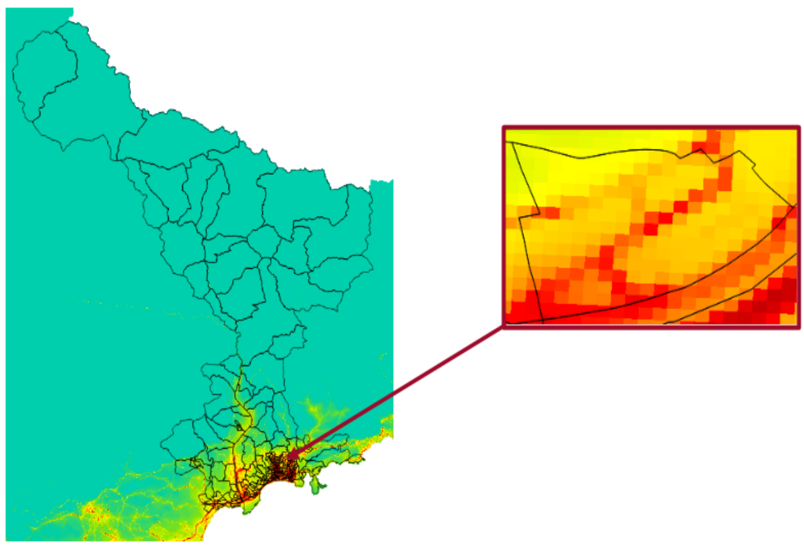

Figure 12 : NO2 concentration on the Nice area

For each IRIS, the objective is to estimate the concentration of the pollutant studied by averaging the values of each 25m square contained in it. For this purpose, a geoprocessing tool provided by the open source GIS software “Qgis” was used, allowing to make statistics for each polygon (our IRIS) from the raster layer data.

This procedure, once carried out with the various Geotiff files, allowed us to obtain an average annual value per IRIS for the pollutants NO_2, PM10 and PM2.5 from 2013 to 2017, which we then linked to socio-economic variables, as well as to the index of socio-economic disfavor.

4.4 First results

These results are based on the internship report written by Anthony Sattler, Master 2 student who spent his six months end-of-study internship at INSEE.

4.4.1 Caracterisation of the index classes

The classification based on the disadvantage index allowed us to identify three classes whose characteristics seem to be very different.

Characteristic	Mean class 1	Mean class 2	Mean class 3	Global mean
Households with 2 cars or more	36.48	11.62	8.18	22.75
Single dwelling	39.86	4.23	2.46	20.40
median standard of living per household	22311.09	18605.50	14451.33	20015.89
principal residences superior to 120m²	12.34	2.66	1.55	7.01
principal residences with a parking or garage	69.42	43.56	34.87	54.79
higher education graduates in the population aged 15 years or over	30.00	26.06[a]	7.89	26.62
workers in the active population	14.11	16.80[a]	33.96	16.74
social housing	3.95	10.19[a]	61.31	10.83
individuals living in single-parent families in the total population	9.87	13.12	21.21	12.19
without diplomas (or very low level) in the population aged 15 or over	23.22	28.85	47.67	27.56
households with more than one person per room	7.95	11.92	26.48	11.10
Unemployed in the active population	10.19	16.27	25.40	14.11
Average amount of housing subsidies per household	273.66	642.65	1585.91	538.42
foreign individuals in the total population	5.06	13.47	27.59	10.59
mean number of individuals per room	0.69	0.79	0.91	0.75
multiple dwellings	53.27	91.29	94.83	74.15
Households who do not own the house	31.96	56.37	78.64	46.73
households without a car	12.18	38.67	43.43	26.88

Table 3 : Values of variables contributing to the disadvantage index for each class

First of all, they are distinguished by very dissimilar financial resources, since the average annual income of individuals living in IRIS belonging to the wealthiest category (class 1) is €22,311, while that of the intermediate class is almost €4,000 less, and only €14,450 for the poorest IRIS class. The amounts received for housing services go in the opposite direction, since they are on average more than 1500€ for the individuals concerned of the IRIS in class 3, 642€ for class 2, to be only 273€ for the first class.

The housing variables indicate that there is a very high share of single-family homes (nearly 40%) in class 1, compared to 4% and 2% for the others, and that conversely, more than 90% of the dwellings in the last two classes are in multiple residences. Class 3 is characterized by a very high share of low-cost housing (61%), much higher than the metropolitan average of about 11%, and less than 4% for class 1. Nearly 80% of individuals living in the most disadvantaged IRIS do not own their own homes, compared to 56% for the middle class and 32% for the privileged class, in which there is a larger share of large housing units (more than 120 m²) up to 12%, about 10 percentage points higher than the others. Finally, while the average number of people per room is concentrated around the average of 0.75 in classes 1 and 2 (respectively 0.69 and 0.79), it is much higher in the third class with 0.91, where more than a quarter of the residences have more than one person per living room.

Secondly, with regard to education and work, the first class has a higher proportion of tertiary graduates (30%) than of non- or low-skilled graduates (23%), while the trend is clearly the opposite in class 3 where there are almost 8% of tertiary graduates compared to 48% without diplomas. The latter has a larger working population, accounting for 34% of the working population, which is twice the average for the metropolitan area, as well as more than a quarter of the unemployed, while the unemployment rate is 16% and 10% in the intermediate and privileged classes.

Finally, first-class households have a larger number of cars, with 36% having two or more, which is only 12% and 8% for the other two categories, with 39% and 43% of households without vehicles respectively. More than two thirds of individuals and IRIS in class 1 have a parking space, a share that drops to 44% for the second and 35% for the third, due to the fact that the majority of IRIS in these classes is located in urban areas. There is a higher proportion of individuals living in single-parent families in class 3 (21%) compared to an average of 12% in the study area.

4.4.2 Cartography of the socio-economic index

The classification of IRIS according to their level of socio-economic disadvantage can first be represented on a map to observe spatial dispersion and identify where the less favoured areas are.

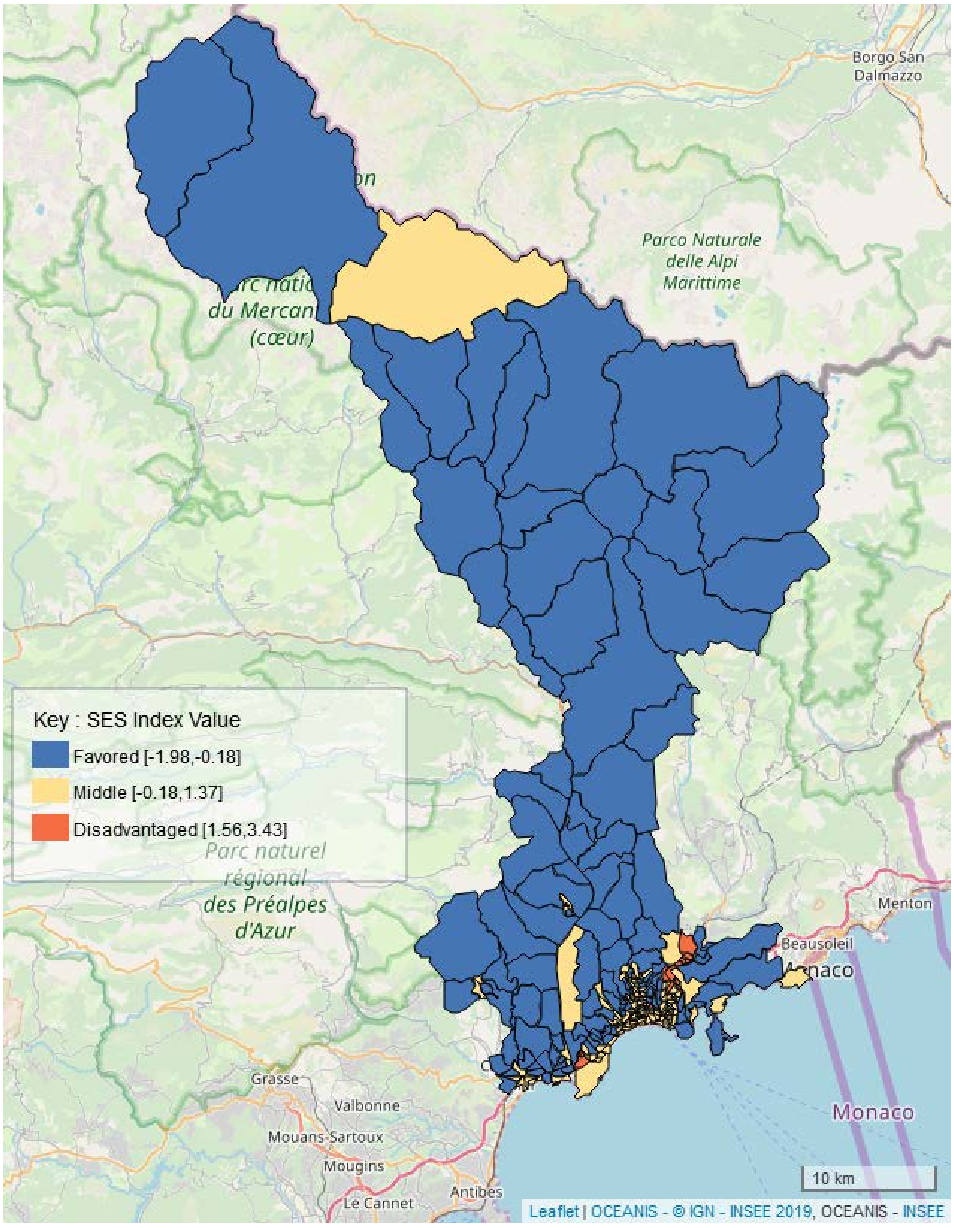

Figure 13 : Socio-economic index value for the Nice metropolis

This map (Figure 6) represents the value of the socio-economic index according to the grouping previously carried out, over the entire Nice Côte d'Azur metropolitan area. We can observe a majority of favoured areas in the northern part of the metropolis. The medium and disadvantaged IRIS seem to be located in the urban heart of the metropolis, i.e. close to the coast as well as in the city of Nice. The majority of the population of the metropolitan area is either in the most favoured class, with 236,126 individuals, or in the intermediate class with 264,649 individuals. The last class has 37,781 individuals.

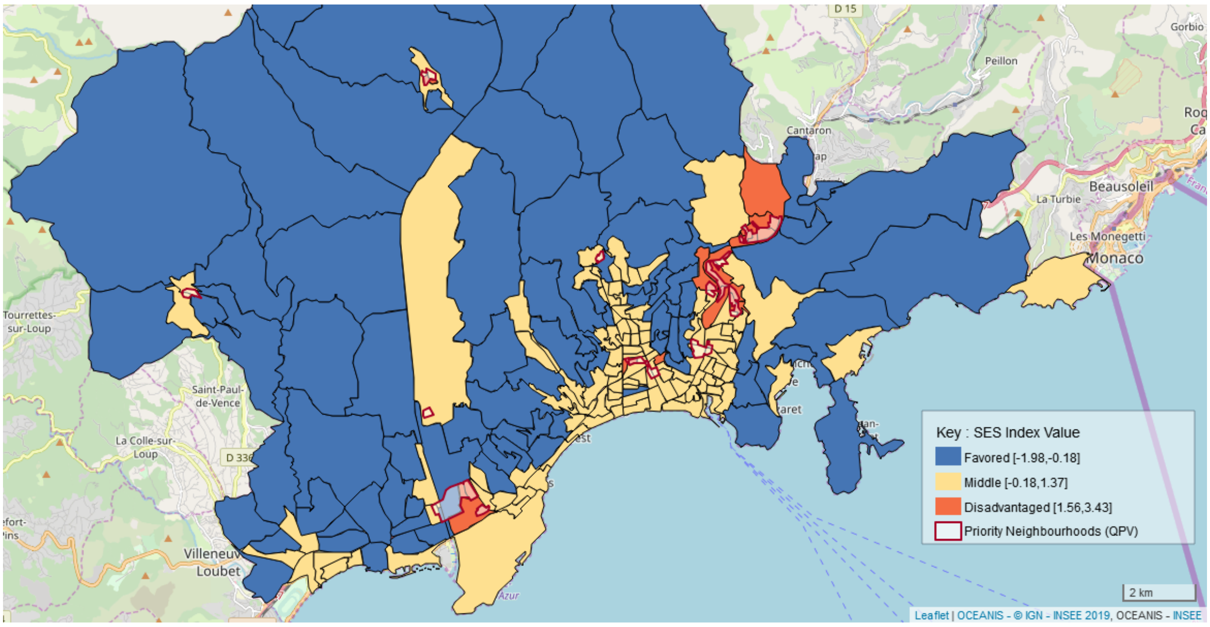

Figure 14 : Socio-economic index on the Nice metropolis - municipalities with IRIS

Let us take a closer look at the IRIS of the commune of Nice and its surroundings, which counts 146 of the 234 studied in total (Figure 7). First of all, it appears that the distribution of socio-economic conditions does not seem to be homogeneous over the city. Indeed, the IRIS identified as favoured (in blue) spread mainly in the northern and western parts of the city, with also a group of three at the southeast end. It should be noted that their surface area seems quite large compared to the others, but the number of inhabitants is only 92,081. In contrast, the medium-level IRIS (in beige), which are mainly concentrated in the southern (coastal) and central-eastern part, comprise the majority of the city's population with 214,701 individuals. Finally, those who are the most disadvantaged, in orange, are found in two particular geographical areas of the city: a main group in the northeast, on a very important road (A8 towards Italy), with a population of 22,783, and a smaller group southwest of Nice, near Saint-Laurent-du-Var, with also a strong presence of the A8 towards Aix-en-Provence, where 8,757 people live. Finally, the last two IRIS, with a population of 4,225, are located in the heart of the city.

4.4.3 Comparison with priority districts of city policy

In order to strengthen the work done, and in particular the detection of areas considered to be the most disadvantaged, it is possible to observe the distribution of priority districts of the city policy (QPV). These neighbourhoods are areas developed by the Commissariat général à l'égalité des territoires (commission for the equality of territories), requiring special intervention by the Ministry of the City. In metropolitan France, they are defined solely according to the poverty concentration criterion defined by INSEE, which uses data computed on squares of 200 metres of sides to compare the income of the inhabitants with the median reference income. The new decree of 30 December 2014 made it possible to define the outlines of these districts for the period 2015 to 2020. They are shown in red on the previous map (Figure 7), most of them in the city of Nice.

The city of Nice has eight qpvs, one of which extends over Saint-Laurent-du-Var and another over Saint-Andre-de-la-Roche. As far as their locations are concerned, the largest qpvs are located in IRIS in the northeast, such as Paillon and Ariane - Le Manoir, as well as southwest of Nice (Les Moulins - Le Point Du Jour, Résidence Sociale Nicéa and Les Sagnes). The three others are located further in the centre of the city are Las Planas, Palais Des Expositions, and Centre. The other two visible on the previous map are the Carros-centre district, in the municipality of Carros to the north of Nice, and Vence-centre to the west of Nice.

The majority of the IRIS where these neighbourhoods are located has been identified as being at a socio-economic disadvantage, and some others are in the middle class. A single qpv in the southwest extends over several IRIS, one of which belongs to the most favoured class. However, most of the population of this district is located in IRIS Les Moulins (in the disadvantaged group).

In the end, it seems that the socio-economic status index is a good approach to identify the most disadvantaged areas, since there is a strong correspondence with the location of priority districts of urban policy, a key element of poverty in an area. In addition, this index does not only depend on the income of the inhabitants, but is a multidimensional indicator that takes into account the different aspects by which social disadvantage and economic poverty can manifest themselves.

4.4.4 Pollution cartography

Following maps (Figure 8, 9) show us the different concentration levels for NO_2 and fine particles 10um in diameter on the IRISes of the municipalities south of the metropolis.

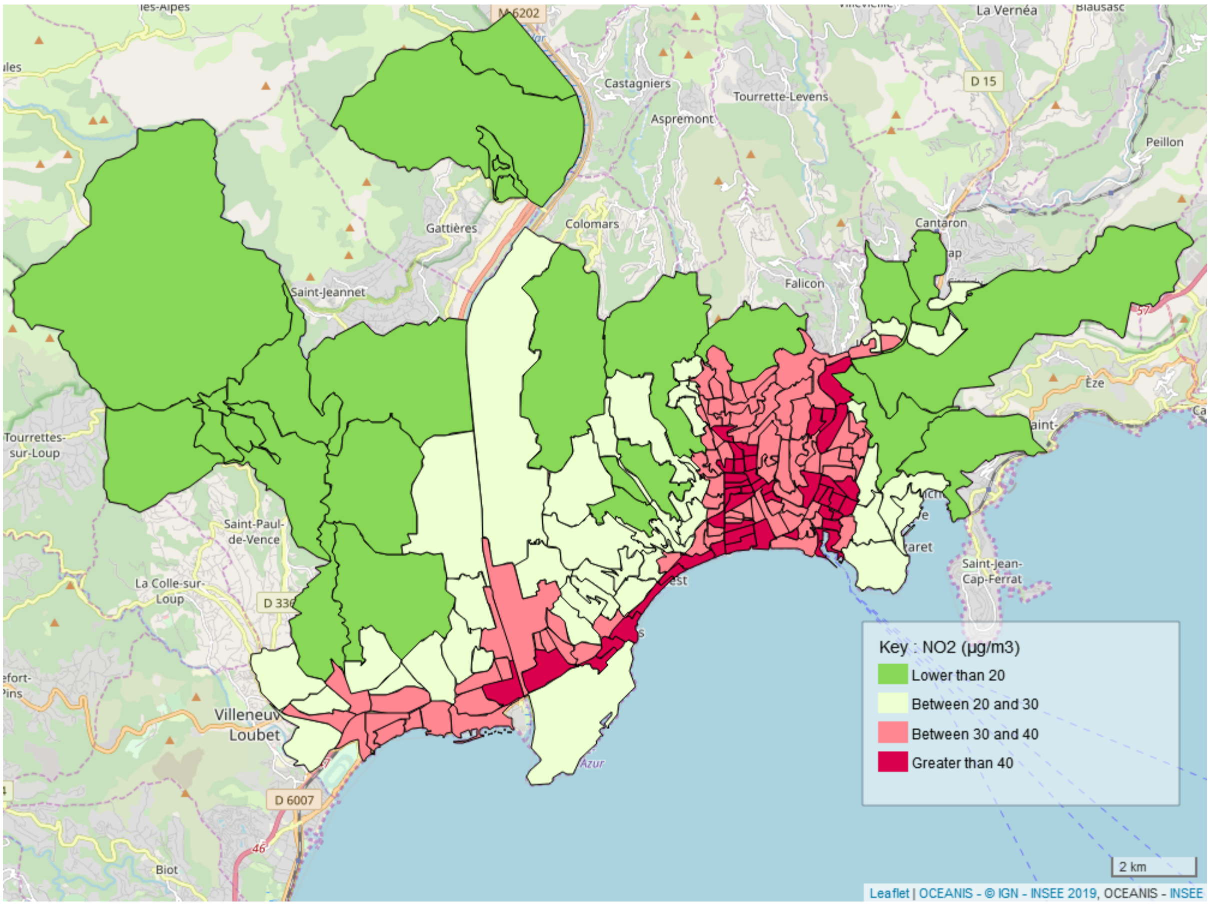

Figure 15 : NO2 concentration on the Nice metropolis

NO_2 is the result of combustion (heating, electricity production, vehicle engines). WHO threshold is a maximum of 40 µg/m^3 in annual mean before health is impacted. Exposure to high levels of NO_2 can lead to a significant inflammation of the respiratory system and to a decrease in lung function. the IRIS that are among the least polluted with NO_2 are located in the peripheral municipalities of Nice, as well as in the northern part of the city, with a concentration of less than 20 µg/m^3, and most of the areas where the concentration is less than 30 µg/ m^3 are located in the west and east of the city. The heart of Nice and the coastline to the west are mainly composed of IRIS where the NO_2 value is between 30 and 40 µg/m^3. Finally, the most polluted IRIS, whose average annual NO_2 concentration exceeds the WHO recommended threshold of 40 µg, are found at the western entrance to the city of Nice, on its coastal part and at its northeast exit. This is the route of the city's main road axis, as detailed in Figure 8.

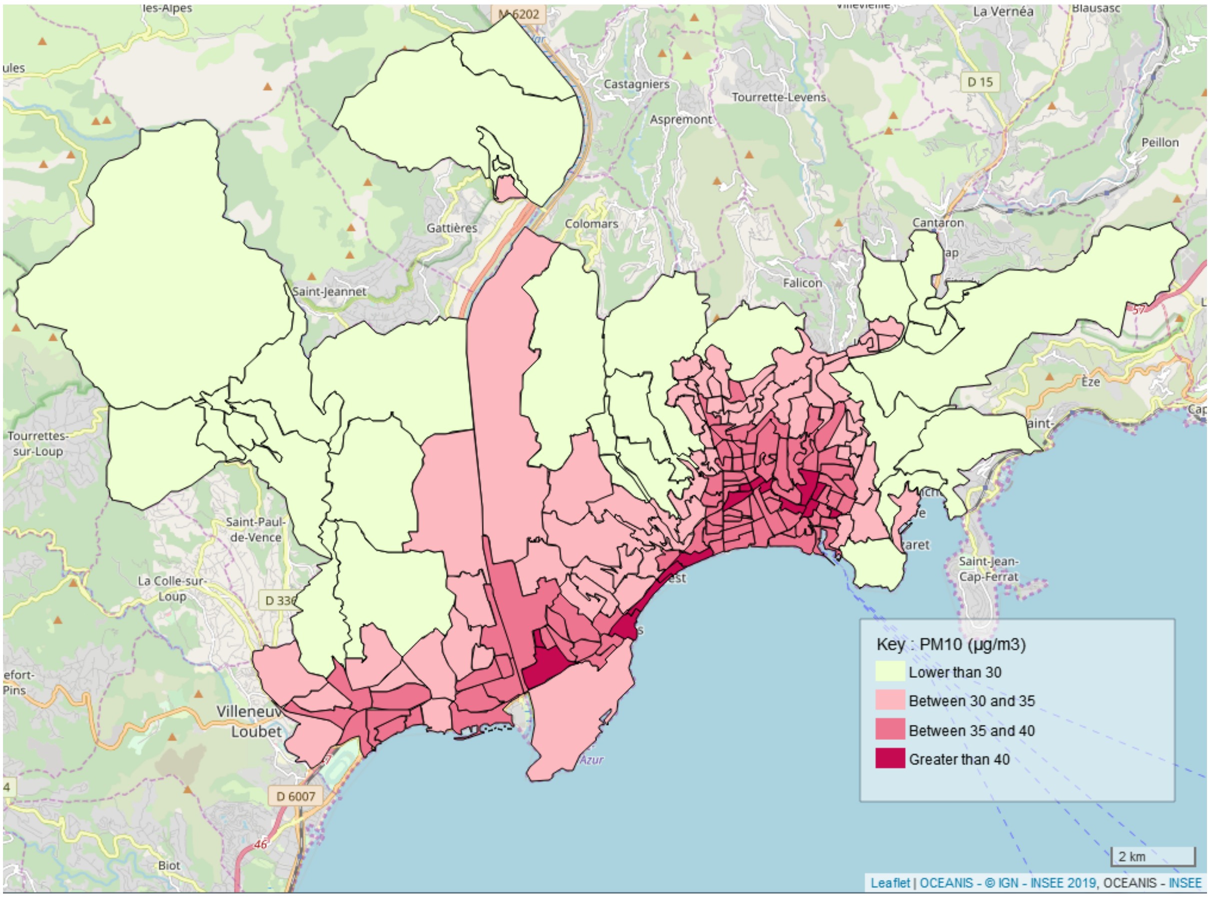

Figure 16 : Mean annual PM10 concentration in the Nice metropolis

Fine particulate matter with a diameter of less than 10 micrometers (PM10) are due to the industrial and transportation sector (mainly diesel combustion). WHO threshold is a maximum of 20 µg/m^3 in annual mean before health is impacted. Exposure to high levels of PM10 can lead to a decrease in life expectancy, cardiovascular disease and lung cancer.

The distribution of PM10 values over municipalities covered by IRIS is quite similar to that of NO_2, and it is the same IRIS that emerge as having the highest pollution concentrations. However, all IRIS exceed the WHO recommended annual average threshold of 20 µg. The most polluted central and coastal IRIS are even exposed to a concentration twice as high as the limit. However, this is not necessarily the most representative of the danger of this pollutant, since there is another average daily threshold not to be exceeded, which is 50 µg.

4.4.5 Exposition to pollution and social vulnerability

In order to highlight the differences in pollution exposure that may exist according to socio-economic category, we carried out ANOVA tests, the principle of which was to determine whether averages are significantly different between groups.

First, let us compare the pollution levels within the three groups previously determined by the hierarchical bottom-up classification.

Mean ± standard deviation	Class 1	Class 2	Class 3
NO_2	24.6 ± 7.8	36.8 ± 7.6	33.7 ± 9.8
PM_{10}	31.8 ± 3.5	36.3 ± 3.2	34.3 ± 4.6
$PM_{2,5}$	12.2 ± 0.9	13.4 ± 0.8	12.8 ± 1.2

Table 4 : Average concentration of pollutants over the period 2013-2017 (µg/m3 ± CI)

Table 4 allows to compare the average concentrations of pollutants over the period 2013-2017 for each group resulting from the classification on the disadvantage index, only on the municipalities with IRIS. Class 1 includes the most favoured IRIS and class 3 the most disadvantaged IRIS.

First of all, for NO_2, it appears that the average concentration in the better-off IRISes is significantly lower than in the other two groups, with 24.6 µg/m³. Indeed, according to the associated non-parametric ANOVA test, the p-value is less than 0.001. The average is 1.5 times the value of central class IRISes, with 36.8 µg/m³, and slightly lower for disadvantaged IRISes (33.7 µg/m³). The average concentrations of the latter two groups are not significantly different, with a p-value of 0.19.

Then, for fine particles PM10, the average is again lower for favoured IRISes, with 31.8 µg/m³. However, the p-values of the Wilcoxon tests indicate that it is significantly different from the intermediate group, with $p<0.001$, but this is not the case with the disadvantaged IRIS group (0.093). The average annual PM10 concentration between the latter two groups is also not significant at 5%, since the p-value of the ANOVA test is 0.068.

Finally, since PM2.5 is closely related to PM10, the results are quite similar. The average of 12.2 µg/m³ for the favoured group is significantly different from those of the intermediate groups with 13.4, but not from that of the disadvantaged, which is 12.8.

The average annual concentrations of NO_2 recorded are, for all classes, below the WHO annual threshold of 40 µg/m³. On the other hand, they are significantly lower for PM10 (threshold at 20 µg/m³) and slightly higher for PM2.5 whose threshold is set at 10 µg/m³. However, as mentioned above, these annual thresholds are given as an indication, the most dangerous for fine particles being exposure for several days consecutive to a high concentration compared to normal (respectively 50 and 25 µg/m³ on average over 24 hours).

For the rest of the descriptive analysis, we will study the evolutions between 2013 and 2017 for the three categories of percentiles at the extreme values of the index, namely below the 20 percentile, between 40 and 60, as well as above the 80 percentile, as was done by the researchers of the Equit'Area project. IRIS below the 20 percentile are the most favoured, while those above the 80 percentile are the most socio-economically disadvantaged. This change allows a better discrimination between results, as IRIS belonging to extreme values of the index present more distinct characteristics.

		<p20		p60_40		>p80	
Mean ± standard deviation		**2013**	**2017**	**2013**	**2017**	**2013**	**2017**
NO_2	µg/m^3	21.7 ± 7.5	21.3 ± 7.2	37.7 ± 5.9	36.4 ± 6.0	38.2 ± 9.6	36.2 ± 9.3
	evolution	-1.8%		-3.4%		-5.2%	
PM_{10}	µg/m^3	34.6 ± 2.9	27.1 ± 3.5	39.8 ± 2.5	33.4 ± 2.9	39.9 ± 3.9	33.7 ± 5.0
	evolution	-21.7%		-16.1%		-15.5%	
$PM_{2,5}$	µg/m^3	12.9 ± 0.8	11.0 ± 0.9	14.3 ± 0.7	12.6 ± 0.8	14.3 ± 1.0	12.7 ± 1.3
	evolution	-14.7%		-11.9%		-11.2%	

Table 5 : Evolution of the average pollutant concentration for the low, intermediate and high categories of the percentiles of the disadvantage index (µg/m3 ± CI)

Comparison of measurements between the 3 categories of percentiles

According to the ANOVA test carried out, the average pollution levels recorded are significantly different between the most favoured category and the other two categories in 2013 and 2017. For both NO_2 and fine particulate matter (PM10 and PM2.5), the value to which populations are exposed is lower for IRIS whose SES index is lower than the 20 percentile (i.e. the most favoured). It then increases sharply for individuals residing in IRIS between the 40th and 60th percentile, as well as those above the 80th percentile, between whom the difference does not appear significant.

Evolution for each category

In general, the average annual concentration of the three pollutants observed tends to decrease from 2013 to 2017.

First of all, we note for NO_2 that its concentration has decreased more rapidly over the period among the most disadvantaged IRIS. Indeed, those above the 80th percentile saw this value fall by more than 5%, while it is 3.4% for the intermediate percentile group, and only 1.8% for the most privileged. The gap between classes narrowed slightly over this period.

On one hand, for PM10, the opposite effect occurs. Over the five years, we observe a very significant decrease - of more than one fifth - in the IRIS below the 20 percentile. On the other hand, it is more moderate for the intermediate group - with a decrease of 16.1% - and for the most socio-economically disadvantaged – with 15.5%.

Finally, the concentration of PM2.5 also has the same trend (since they are highly related to PM10) but with a lesser effect. We notice a drop in PM2.5 of a little less than 15% for the wealthiest, while it is only 11.9% for the 40 to 60 percentile, and finally 11.2% for the IRIS above the 80 percentile.

4.5 Follow-up studies

4.5.1 Using NSIs data to improve the estimation of exposure to pollution

Our first axis of improvement for future studies is to improve the quality of the estimation of people's exposure to air pollution.

Even if we keep using annual averages per pollutant from the Atmosud modelling, it is possible to calculate an average concentration at IRIS differently. Indeed, since INSEE has the geolocation of individuals via the Fidéli database, we could take only into account the areas where there are inhabitants and no longer the entire surface area of the IRIS.

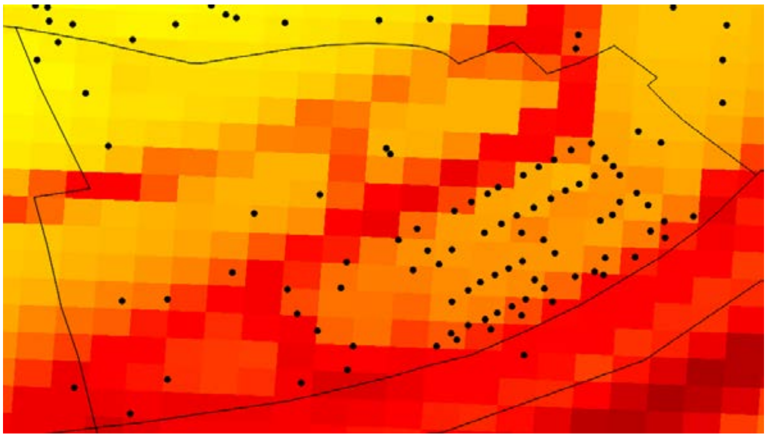

Figure 17 : Superposition of points (individuals) on vectorial (Iris) and raster layers (pollutant modeling)

It is thus possible to consider only the 25m squares in which the dwellings are included to determine a value of pollution at the IRIS. It may also be considered to weight each value by the number of

individuals it concerns, in order to produce a number corresponding to an exposure to pollution taking into account the number of people affected.
In addition, a disadvantage index can be calculated for a period before 2015, for example using the 2010 census. This would allow us to see the evolution of the disadvantage over the two periods as well as the differences in population exposure between the two periods.
Then, rather than working on annual averages as is the case for this first approach, a study of pollution at a finer time scale may reveal a more real exposure to which individuals are confronted. Indeed, the concentration of air pollution in a place is far from constant over time, and varies greatly depending on traffic and weather in particular. Here is a first diagram (Figure 11) that explains the problem more closely.

Figure 18 NO2 measurements by two sensors for one week

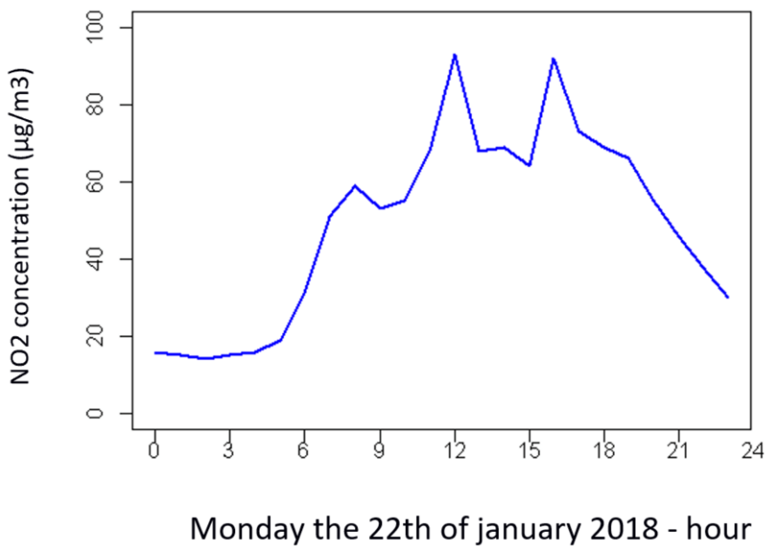

Figure 19 Nitrogen dioxyde measurements by one sensor for one day - sensor at the Nice English walking

Figure 10 indicates the NO_2 concentration per day over a week for the sensors located at Nice Airport and Promenade des Anglais (English walkway). It should be noted that there is a third fixed station, which is located in the Arson district (north of the port). We can see a fairly significant variation depending on the day of the week, regardless of the location. At the airport, the NO_2 concentration drops from 12 µg/m³ on Sunday to a value three times higher on Monday (37 µg/m³), and which tends to decrease throughout the week. At the walkway, the observed values are higher, and the difference between the highest and lowest day is also about 20 µg/m³ with a slight increase on Friday and Saturday compared to Thursday over that week. The variability at one point thus seems quite high, but the same is true for a day.

By observing in more detail the day of Monday at the sensor located at the promenade, we can see a strong variation according to the hours of the day. After a first peak at 8am with 60 µg/m³, the NO_2 concentration reaches its maximum around 12pm and 4pm reaching 90 µg/m³. In this example, air pollution can be a health hazard at the observed times of day. However, a large proportion of the individuals exposed to these high pollution episodes do not correspond to the resident population (i.e. identified at their place of residence) as used in this report, but to the present population (where the individuals are during the day). The resident population approach is certainly justified, since it takes into account all individuals who are at their place of residence during the day as well as the exposure suffered by the housing in which they live, but it would need to be supplemented by resident population data. To date, there are relatively few elements to carry out such an approach, and INSEE is working on projects that could lead to an estimate of the population as it is during the day. However, this remains complicated to determine and requires further knowledge. Other elements can also be used to approach the notion of present population, such as the Permanent Base of Equipment (BPE), which can be an added value in terms of places where some of the individuals during the day can be found (schools, retirement homes, museums, etc.).

In addition, this partnership is an opportunity to build on each other's strengths. Thus, in order to focus on a finer time frame in terms of pollution surveys, INSEE and Imredd will be able to agree on the installation of temporary sensors in strategic areas, for example those identified as the most disadvantaged by the disadvantage index, but also in others for more precise comparisons. Indeed, Imredd, which works with the metropolis of Nice in the management of the Smart city, can intervene in the implementation of these devices.

4.5.2 Understanding the reasons for pollution, thanks to the smart city mobility data

In the medium term, notions of mobility and vehicle flows will be included in this study in order to explain the causes of pollution, and therefore to act against these causes. This is one of the areas of work of Imredd, which has data related to mobility and produces adapted models (multi-agent simulations). The interest in relating air pollution, socio-economic characteristics and mobility data is clearly visible in figure 12 linking the socio-economic classes of the IRIS with the main roads.

We clearly recognize the points raised when observing the locations of the disadvantaged IRIS, namely that the two most underprivileged areas are located at the city's entrance and exit gates, to the southwest and northeast. Apart from the A8 motorway, which allows the city to be bypassed by less populated areas, but on a tollgate section, which may discourage some motorists, the main road starts at the Saint-Laurent-du-Var exit and continues along the coast before going up to the motorway and the east exit of the city. Understanding the issue of mobility can thus make it possible to better understand travel in the city, and subsequently implement measures to move a certain proportion of traffic, all with the aim of reducing the social inequalities in health to which residents of more disadvantaged and pollution-prone areas may be exposed.

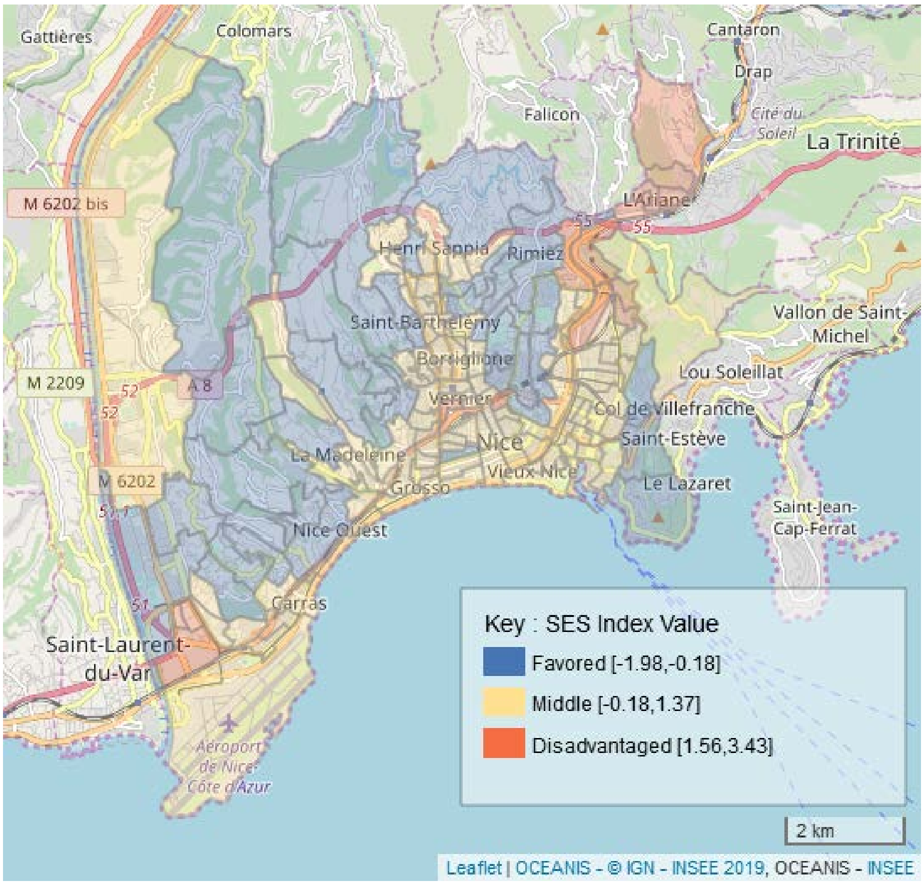

Figure 20 : Socio-economic index and transportation axis

Finally, still in relation to mobility, the long-term pursuits will be to assess the impact of the opening of tram lines in cities, in order to determine whether this has had an influence on the distribution of the population in terms of socio-economic characteristics. Indeed, real estate prices tend to increase with the improvement of the quality of life in urban areas, and disadvantaged populations are therefore forced to move to less pleasant and more polluted neighbourhoods.

5 The ISTAT's project in detail

5.1 Partnership

ISTAT is working to establish a partnership with "ARPA Lazio'' that is the Regional Agency for Environmental Protection to which the city of Rome belongs. A first meeting has taken place in May 2019, and a mutual interest to work together on this case study has emerged. The common aim is to produce experimental statistics that links the ISTAT Census data on population and dwellings with the air pollution data gathered by the Agency, through their smart devices. The next aim is to formalize the agreement between ISTAT and ARPA Lazio so that the collaboration could be started.

5.2 Data

<u>Social data (population and dwellings)</u>

The social data needed to carry on the case study are the data on population and dwellings collected by the last Census. In the following table all the useful variables are listed. The variables have been grouped by the following areas of interest: family, immigration, education, employment, population and housing.

Here is the list of the 116 variables that have been determined at the census section

Area	Variable (in number)	Source
Family and household	Families in rental accommodation Families in owned housing Resident families - total Resident families - 1 component Resident families - 2 component Resident families - 3 component Resident families - 4 component Resident families - 5 component Resident families - 6 component and over	Census 2011 - population
Immigration	Foreigners and stateless persons residing in Italy - total Foreigners and stateless persons residing in Italy - males Foreigners and stateless persons residing in Italy - age 0-29 years Foreigners and stateless persons residing in Italy - age 30-54 years Foreigners and stateless persons residing in Italy - aged over 54 years Foreigners and stateless persons residing in Italy - males - age 0-29 years Foreigners and stateless persons residing in Italy - males - age 30-54 years Foreigners and stateless persons residing in Italy - males - aged over 54 years Foreigners persons residing in Italy - from Europe Foreigners and stateless persons residing in Italy - from Europe Foreigners and stateless persons residing in Italy - from Africa Foreigners and stateless persons residing in Italy - from America Foreigners and stateless persons residing in Italy - from Asia Stateless persons residing in Italy Foreigners persons residing in Italy	Census 2011 - population
Education	Resident population - university graduate Resident population - high school graduate Resident population - middle school graduate Resident population - elementary school graduate Resident population - literate males Resident population - illiterate males Resident population - males aged six years and over Resident population - males - university graduate Resident population - males - high school graduate Resident population - males - middle school graduate Resident population - males - elementary school graduate Resident population - literate males Resident population - illiterate males	Census 2011 - population
Employment	Resident population - total - 15 years and over - Labour Force Resident population - total - 15 years and over - Employed Resident population - total - 15 years and over - Unmployed Resident population - males - 15 years and over - Labour Force Resident population - males - 15 years and over - Employed Resident population - males - 15 years and over - Unmployed	Census 2011 - population

	Resident population - total - 15 years and over - Not Labour Force Resident population - males - 15 years and over - Not Labour Force Resident population - total - 15 years and over - Housewife Resident population - total - 15 years and over - Students	
Population	Resident population - total Resident population - males Resident population - females Resident population - unmarried Resident population - married Resident population - widowers Resident population - males - divorced Resident population - males - unmarried Resident population - males - married Resident population - widowers Resident population - divorced Resident population - age < 5 years Resident population - age 5 - 9 years Resident population - age 10 - 14 years Resident population - age 15 -19 years Resident population - age 20 - 24 years Resident population - age 25 -29 years Resident population - age 30 - 34 years Resident population - age 35 -39 years Resident population - age 40 - 44 years Resident population - age 45 -49 years Resident population - age 50 - 54 years Resident population - age 55 -59 years Resident population - age 60 - 64 years Resident population - age 65 -69 years Resident population - age 70 - 74 years Resident population - age > 74 years Resident population - males - age < 5 years Resident population - males - age 5 - males - 9 years Resident population - males - age 10 - 14 years Resident population - males - age 15 -19 years Resident population - males - age 20 - 24 years Resident population - males - age 25 -29 years Resident population - males - age 30 - 34 years Resident population - males - age 35 -39 years Resident population - males - age 40 - 44 years Resident population - males - age 45 -49 years Resident population - males - age 50 - 54 years Resident population - males - age 55 -59 years Resident population - males - age 60 - 64 years Resident population - males - age 65 -69 years Resident population - males - age 70 - 74 years Resident population - males - age > 74 years	Census 2011 - population
Housing	Buildings - total Residential buildings Commercial buildings Residential buildings built before 1919 Residential buildings built between 1919 and 1945 Residential buildings built between 1946 and 1960 Residential buildings built between 1961 and 1970 Residential buildings built between 1971 and 1980 Residential buildings built between 1981 and 1990 Residential buildings built between 1991 and 2000 Residential buildings built between 2001 and 2005 Residential buildings built after 2005 Residential buildings with one floor Residential buildings with two floors Residential buildings with three floors Residential buildings with for floors or more Residential buildings with one apartments Residential buildings with two apartments Residential buildings with number of apartments between three and four Residential buildings with number of apartments between five and eight Residential buildings with number of apartments between nine and fifteen Residential buildings with sixteen apartments or more Total number of apartments in residential buildings Buildings for residential use with excellent state of preservation Buildings for residential use with good state of preservation Buildings for residential use with no good state of preservation Buildings for residential use with poor state of preservation	Census 2011 - dwellings

Table 6 : Variables determined at the census section, conerning population and dwellings

Pollution data

Pollution data come from the regional agency for monitoring air quality in the Lazio country, called ARPA Lazio. The data are available on its web site[20] for each main city of Lazio region and they can be downloaded in .txt and .pdf format. The pollution data are expressed in micrograms per cubic metre (µg/m3) and they are referred at the main air pollutants, which are sulfur dioxide (SO_2), NO_2, carbon monoxide (CO), fine particulate matter (PM10 and PM2.5), ozone (O3) and Benzene. The data are available at daily, monthly and yearly level, for all days between 1999 and now, but from a geographical point of view they refer to each official station and are collected in thirteen official station disseminated in the city of Rome. More detailed data, at geographic level, are published on the ARPA's web site only through a map that is the result of a very complex mathematical model illustrated in the next paragraph. The spatial scale of this model corresponds to squares of 1 Km side, except countryside areas where the spatial scale is only 4 Km².

With regard to the methods and relevant aspects identified for the study, we have done the following preliminary considerations:

- both data-sources, census data and air pollution data, are geo-localized
- the data sources can be linked at census section level
- during data analysis, we chose to consider also the temporary population[21]. This is an improvement addition output for our study.
- for the population we have decided to use the following census variables: civil status, age classes of residential population, level of education, employment status; regarding dwellings, the variables we would include in our study are: number of rooms, date of construction, kind of dwelling; the use of other variables will be evaluated. The reason for this chose is that all these variables are useful for having relevant profiles on pollution impact.
- regarding fiscal data we have considered to use the level of education as a proxy of the income
- regarding the possibility to also use citizen science data[22] on pollution, we think that it could be a very interesting data source, even if they have some calibration problems that make the data gathered not very reliable. The sensors used by the citizens to collect air pollution data cost a thousand times less than the official sensors used by the environmental protection agencies, so that, we understand well that the quality of the data collected is not comparable. We had thought of using them anyway, simply as an additional data source, but for the city of Rome, chosen as pilot city, this kind of data is not available. For all these reasons, we have decided not to use air pollution data collected by citizens.

Mathematical model for extending data

In the city of Rome there are only thirteen official sensors, but ARPA Lazio is able to produce very accurate estimated values of pollution, due to a very sophisticated modeling system, developed by the Agency. The modeling system, used to predict a near-real time reconstruction of air quality, uses the most up-to-date techniques for: (i) description of emissions (ii) transport and chemical

[20] http://www.arpalazio.net/main/aria/sci/

[21] temporary population: people who do not usually live in the house: guests, occasionally living, temporarily present there but resident elsewhere

[22] we refer to air pollution data collected by citizens through the use of small sensors, in a context of open source projects

transformation of pollutants into the atmosphere. The main inputs used by the model are: (i) an Eulerian model for the dispersion and chemical reactions of pollutants in the atmosphere (FARM), (ii) a prognostic meteorological model for downscaling weather forecasts from the synoptic scale to the local scale (RAMS). The system is able to make predictions on the concentration of the most interesting pollutants: sulfur dioxide (SO_2), NO_2, CO, fine particulate matter (PM10), ozone (O3) and Benzene. The estimated data of the model are mapped on a grid of 1 km per 1 km. A complete documentation on the methodology adopted by the modeling system is published on the ARPA Lazio's web site[23]. The output data of the model are not available on the web site; but it is possible to visualize the geographical chromatic map in near-real time reconstruction as shown in the following figure 14.

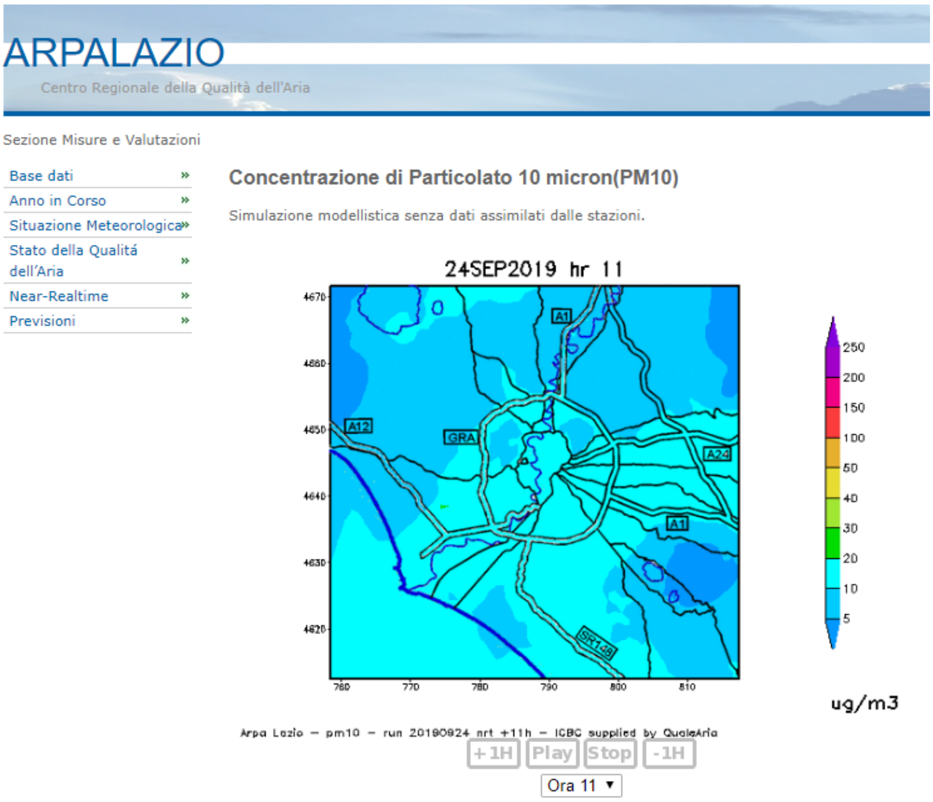

Figure 21: Example of near real-time, geographical chromatic map in the city of Rome (source: ARPA Lazio web site)

5.3 Methodology

The census data are available at different aggregation levels; the most detailed level is the census section. For urban areas, with high population density, the size of the census section is comparable with the grid (1 km square) on which the air pollution data are mapped. In the city of Rome there are about 13,000 census sections and the average area is less than $1km^2$.

Both the census data and the air pollution data are geo-localized; with the GIS (Geographic information system) methodology, we can link the two data sources reporting them at the census section level. In this way, we can elaborate the average data of the pollution and the social profiles of the population and referring them to a single territorial dimension. This elaboration allows us to study the average population profiles with a high risk of exposure to air pollution in the territory of Rome.

[23] http://www.arpalazio.net/main/aria/doc/SM/previsioni.php

5.4 First results

At the beginning of our work on this case study we had only a rough idea on how to study the socio-economic characteristics of people exposed to pollution with the aid of smart sensors and of official data about dwellings. In these months we have worked on the following tasks: (i) to find a pilot city, equipped with smart sensors on which it was possible to access their air pollution data; (ii) to identify the partner to involve in the project; (iii) to analyze the data gathered by the sensors; (iv) to analyze the census data and identify the variables of population and dwellings useful for the study; (v) define the methodology to apply at the case study. All these aspects have been described in the previous paragraphs and represent a first preliminary result. The contribution of ISTAT is therefore a feasibility study including data analysis and definition of the methods to apply. However, the study has still to be implemented.

5.5 Follow-up studies

Considering that: (i) the work package L is only a M12 project and has been defined as a preparation WP, as mentioned in its title «Preparing intelligent statistics», (ii) ISTAT started the WP without any partnership or collaboration; the only achievable goal was to put the ground for the implementation of a future experimental statistics. In addition, ISTAT is going to continue defining the partnership with the Regional Environmental Protection Agency; this point is certainly essential for a follow-up study.

6 Similarities and differences

The comparison of the two studies cannot be complete because they are in two different stages of progress. INSEE started before and it has established relationships with IMREDD (Mediterranean Institute for Risks, Environment and Sustainable Development) so that it has completed a first pilot project. ISTAT, instead, started later and it is still in stage of defining the partnership with the Agency for Environmental Protection.

Nevertheless, we can make a first comparison between the INSEE's case study and ISTAT's feasibility study, from a very general position; from this comparison emerge the following similarities and differences:

Similarities:

- the objective to be achieved is the same: to study the socio-economic characteristics of the population exposed to air pollution
- both INSEE and ISTAT use the population and housing data deriving from the respective censuses
- the variables relating to pollution concern the same pollutants and are expressed in the same unit of measurement ($\mu g/ m^3$)
- the territorial level of reference is very detailed: in both the studies is an inframunicipal statistical breakdown produced by the respective INSs (IRIS for INSEE and Census Section for ISTAT)
- methodology adopted: like INSEE, ISTAT is going to use GIS methodologies to link the several data sources related to different geographical areas.

Differences:

- the number of pollution sensors, considering the surface of the pilot city, is greater in Nice than in Rome (3 fixed + 5 mobile in Nice for a surface of 72 km^2; 13 fixed in Rome for a surface of 1285 km^2)
- the spatial scale of the model to estimate pollution in Rome is a 1 Km square, while in Nice for all the pollutants except the ozone is only 25 m square

Insee also used geolocalized fiscal data to characterize people's socio-economic status.

7 References

ARPA Lazio, “Sistema modellistico per la qualità dell'aria”, http://www.arpalazio.net/main/aria/doc/SM/previsioni.php

European Commission. 2016. « Paris agreement». 23 novembre 2016. https://ec.europa.eu/clima/policies/international/negotiations/paris_fr.

Deguen, Séverine. 2013. « Exposition à la pollution atmosphérique et inégalités sociales de santé », 3.

DGFiP, Traitement Apur, Mise à disposition Ville de Paris (DU/STDF), DRIEA, et CG93. 2016. « Parcelle Cadastrale ».

Ecole des Hautes Etudes en Santé Publique. 2008. « Equit’Area - Accueil ». 2008. https://www.equitarea.org/index.php/fr/.

European Environment Agency. 2018. « Unequal Exposure and Unequal Impacts : Social Vulnerability to Air Pollution, Noise and Extreme Temperatures in Europe », n° 22: 102.

Faburel, Guillaume. 2008. « Les inégalités environnementales comme inégalités de moyens des habitants et des acteurs territoriaux : Pour que l’environnement soit un facteur réel de cohésion urbaine ». *Espace populations sociétés*, n° 2008/1 (juin): 111-26. https://doi.org/10.4000/eps.2430.

Genes. s. d. « Accueil ». groupe-genes. Consulté le 11 septembre 2019. http://www.groupe-genes.fr/index.php.

Imredd. 2019. « IMREDD - Institut d’Innovation et de Partenariats de Université Côte d’Azur ». 2019. https://imredd.fr/#qui-sommes-nous.

Inpes (Institut national de prévention et d’éducation pour la santé), Louis Potvin, Catherine M Jones, et Marie-Josée Moquet. 2012. *Réduire les inégalités sociales en santé*. Saint-Denis (Paris): Institut national de prévention et d’éducation pour la santé.

Khreis, Haneen, Anthony D. May, et Mark J. Nieuwenhuijsen. 2017. « Health Impacts of Urban Transport Policy Measures: A Guidance Note for Practice ». *Journal of Transport & Health* 6 (septembre): 209-27. https://doi.org/10.1016/j.jth.2017.06.003.

Kihal, Wahida, Cindy Padilla, et Séverine Deguen. 2015. « INTERVENTION TERRITORIALE VISANT A LUTTER CONTRE LA POLLUTION ATMOSPHERIQUE ET EQUITE EN MATIERE DE SANTE », 18.

Lalloué, Benoît, Séverine Deguen, Jean-Marie Monnez, Cindy Padilla, Wahida Kihal, Denis Zmirou-Navier, et Nolwenn Le Meur. 2015. « SesIndexCreatoR: An R Package for Socioeconomic Indexs Computation and Visualization ». *Open Journal of Statistics* 05 (04): 291-302. https://doi.org/10.4236/ojs.2015.54031.

Lalloué, Benoît, Jean-Marie Monnez, Cindy Padilla, Wahida Kihal, Nolwenn Le Meur, Denis Zmirou-Navier, et Séverine Deguen. 2013. « A Statistical Procedure to Create a Neighborhood Socioeconomic Index for Health Inequalities Analysis ». *International Journal for Equity in Health* 12 (1): 21. https://doi.org/10.1186/1475-9276-12-21.

Lalloué, Benoît, Jean-Marie Monnez, Cindy Padilla, Wahida Kihal, Denis Zmirou-Navier, et Séverine Deguen. 2015. « Data Analysis Techniques: A Tool for Cumulative Exposure Assessment ». *Journal of Exposure Science & Environmental Epidemiology* 25 (2): 222-30. https://doi.org/10.1038/jes.2014.66.

Larmarange, Joseph. 2018. « Multicolinéarité dans la régression ». 2018. http://larmarange.github.io/analyse-R/multicolinearite.html.

Métropole Nice Côte d'Azur. 2014. « Contrat de ville ».

OpenDataSoft. 2016. « C'est quoi la Smart City ? Une introduction à la ville intelligente ». opendatasoft. 29 avril 2016. https://www.opendatasoft.com/fr/blog/2016/04/29/cest-quoi-la-smart-city-une-introduction-a-la-ville-intelligente.

Padilla, Cindy, Wahida Kihal-Talantikit, Sandra Perez, et Severine Deguen. 2016. « Use of Geographic Indicators of Healthcare, Environment and Socioeconomic Factors to Characterize Environmental Health Disparities ». *Environmental Health* 15 (1): 79. https://doi.org/10.1186/s12940-016-0163-7.

Padilla, Cindy, Wahida Kihal-Talantikite, Verónica M. Vieira, Philippe Rossello, Geraldine Le Nir, Denis Zmirou-Navier, et Severine Deguen. 2014. « Air Quality and Social Deprivation in Four French Metropolitan Areas—A Localized Spatio-Temporal Environmental Inequality Analysis ». *Environmental Research* 134 (octobre): 315-24. https://doi.org/10.1016/j.envres.2014.07.017.

Stéfan Lollivier. s. d. « Description Fideli ».