# D3.1 Urban ReLeaf data ecosystem map



Funded by the European Union

UK Research and Innovation

## Document Factsheet

| | |
|---|---|
| **Grant agreement** | 101086638 |
| **Call identifier** | Horizon-CL6-2022-Goverance-01-08 |
| **Project title** | Urban ReLeaf: Citizen-powered data ecosystems for inclusive and green urban transitions |
| **Project start date** | 01/01/2023 |
| **Duration** | 48 months |
| **Instrument** | Horizon Europe |
| **Type of action** | Innovation action |
| **Deliverable number** | D3.1 |
| **Deliverable name** | Urban ReLeaf data ecosystem map |
| **Work package** | WP3: BUILD: BRIDGE CITIZEN-POWERED AND AUTHORITATIVE DATA ECOSYSTEMS |
| **Task** | T3.1 Map urban greening data ecosystems and identify opportunities for citizen-powered contributions |
| **Deliverable lead** | ICCS |
| **Deliverable type** | Report |
| **Due date** | 30/06/2023 |
| **Review completed** | 30/06/2023 |
| **Submission date** | 30/06/2023 |
| **Document version** | 1.0 |



## Disclaimer



## Copyright notice



## Acknowledgement

Funded by the European Union

UK Research and Innovation

## Reference

Please cite this work as:

Karagiannopoulou, Katerina. (2023). Urban ReLeaf data ecosystem map. D3.1 for the Horizon Europe and Innovate UK co-funded project Urban ReLeaf under grant agreement ID: 101086638. 10061290 and 1004179 DOI: 10.5281/zenodo.8085391

| | |
|---|---|
| **Deliverable name** | Urban ReLeaf data ecosystem map |
| **Authors** | Katerina Karagiannopoulou (ICCS) |
| **Contributors** | N/A |

| | |
|---|---|
| **Short description** | Deliverable 3.1 presents the data ecosystem mapping for all the Urban ReLeaf pilot cities using the Open Data Institute's Data Landscape Playbook methodology. |
| **Keywords** | data ecosystems, mapping, open data, data landscape |

| **Dissemination level** | | |
|---|---|---|
| **PU** | Public | X |
| **SEN** | Sensitive – limited under the conditions of the GA | |
| **EU-R** | Classified information - Restricted under EC Decision 2005/444 | |
| **EU-C** | Classified information - Confidential under EC Decision 2005/444 | |

**History**

| Version | Date | Released by | Comment |
|---|---|---|---|
| 0.1 | 04.04.2023 | Katerina Karagiannopoulou (ICCS) | Document preliminary preparation. Creation of the table of contents |
| 0.5 | 09.06.2023 | Katerina Karagiannopoulou (ICCS) | The updated version of the document |
| 0.6-0.9 | 30.06.2023 | Katerina Karagiannopoulou (ICCS), Athanasia Tsertou (CIBOS), Inian Moorthy (IIASA) | Modifications according to reviewers' feedback; Appendix B added |
| 1.0 | 30.06.2023 | Katerina Karagiannopoulou (ICCS) | Final version for submission |



**Project Partners**

# Table of Contents

## Acronyms

| | |
|---|---|
| **3D** | 3-Dimensional |
| **AGB** | Above-Ground Biomass |
| **AI** | Artificial Intelligence |
| **API** | Application Programming Interface |
| **AQ** | Air Quality |
| **AQI** | Air Quality Index |
| **AQMA** | Air Quality Monitoring Area |
| **ArbPIX** | Arnold Arboretum Plant Image |
| **ASCII** | American Standard Code for Information Interchange |
| **AWS** | Amazon Web Services |
| **BC** | Black Carbon |
| **CC** | Creative Commons |
| **CIMLK** | Central Instrument of Monitoring the Air Quality |
| **CKAN** | Comprehensive Knowledge Archive Network |
| **CNN** | Convolutional Neural Network |
| **CO2** | Carbon Dioxide |
| **CS** | Citizen Science |
| **CSV** | Comma-Separated Values |
| **DBH** | Diameter Breast Height |
| **DEC** | Data Ethics Canvas |
| **DEFRA** | National Department for Environment Food and Rural Affairs |
| **DEM** | Data Ecosystem Mapping |
| **DIY** | Do-It-Yourself |
| **DLP** | Data Landscape Playbook |
| **EC** | European Commission |
| **EEA** | European Environmental Agency |
| **EEA** | European Environmental Agency |
| **EFISCEN** | European Forest Information Scenario |
| **EMAC** | Directorate-General for Territory and The Municipal Environment Company of Cascais |
| **EO** | Earth Observation |
| **EPD** | Edible Plant Database |
| **ESRI** | Environmental Systems Research Institute |
| **EU** | European Union |
| **EUPL** | European Union Public Licence |
| **FAIR** | Findable, Accessible, Interoperable, Reusable |
| **FISE** | Forest Information System for Europe |
| **ForestGEO** | Forest Global Earth Observatory |
| **FTP** | File Transfer Protocol |
| **GDA** | Governmental Data Assets |
| **GDB** | Geospatial Database |
| **GEO** | Group of Earth Observations |
| **GEO-BON** | Geo Biodiversity Observation Network |
| **GEO-CITSCI** | Geo Citizen Science |
| **GeoJSON** | Geospatial JavaScript Object Notation |
| **GeoTIFF** | Geospatial Tag Image File Format |
| **GHG** | Greenhouse Gases |
| **GIS** | Geographic Information System |
| **GPS** | Global Positioning System |
| **In/Ex** | Inclusion and Exclusion |
| **IPCC** | International Panel on Climate Change |
| **JPEG** | Joint Photographic Experts Group |

| | |
|---|---|
| **JSON** | JavaScript Object Notation |
| **KML** | Keyhole Markup Language |
| **KNMI** | Royal Netherlands Meteorological Institute |
| **KToIs** | Key Terms of Interest |
| **LCP** | Large Combustion Plant |
| **LPA** | Local Planning Authorities |
| **LULC** | Land Use/Land Cover |
| **MIT** | Massachusetts Institute of Technology |
| **NAPMN** | National Air Pollution Monitoring Network |
| **NBN** | National Biodiversity Network |
| **NHS** | Natural Heritage Site |
| **NO** | Nitrogen Oxide |
| **NO2** | Nitrogen Dioxide |
| **NVC** | National Vegetation Classification |
| **O3** | Ozone |
| **OD** | Open Data |
| **ODbL** | Open Database Licence |
| **ODE** | Open Data Ecosystem |
| **ODGE** | Open Data "Governmental" Ecosystem |
| **ODI** | Open Data Institute |
| **OGC** | Open Geospatial Consortium |
| **OGL** | Open Government Licence |
| **OKF** | Open Knowledge Foundation |
| **OSM** | OpenStreetMap |
| **PANACEA** | Panhellenic Infrastructure for Atmospheric Composition |
| **PDF** | Adobe Reader Portable Document Format |
| **PECAC** | Strategic Plan for Climate Change Adaptation Action |
| **PM** | Particulate Matter |
| **PNG** | Portable Network Graphics |
| **QAQC** | Quality Assurance, Quality Control |
| **QART** | Monitoring and Urban Environmental Mapping Solutions |
| **QoL** | Quality of Life |
| **R&D** | Research and Development |
| **RC100** | 100 Resilient Cities |
| **RCP** | Representative Concentration Pathway |
| **RIVM** | National Institute of Health And Environment |
| **SDI** | Spatial Data Infrastructure |
| **SEPA** | Scottish Environment Protection Agency |
| **SIMD** | Scottish Index of Multiple Deprivation |
| **SO2** | Sulphur Dioxide |
| **SPA** | Special Protection Area |
| **SW** | Software |
| **TOC** | Topsoil Organic Carbon Content |
| **TPO** | Tree Preservation Orders |
| **UFP** | Ultrafine Particles |
| **UHI** | Urban Heat Island |
| **UK** | United Kingdom |
| **UNECE** | United Nations Economic Commission for Europe |
| **UN-SDGs** | United-Nation Sustainable Development Goals |
| **UV** | Ultraviolet |
| **VHR** | Very-High-Resolution |
| **VOC** | Volatile Organic Compounds |
| **VSIA** | Latvian State Roads |
| **WDL** | World Data League |
| **WFS** | Web Feature Service |

| | |
|---|---|
| **WMS** | Web Map Service |
| **WMTS** | Web Map Tile Service |
| **WP** | Work Package |
| **YPEN** | Hellenic Ministry of Environment and Energy |

## List of Figures

## List of Tables

# Executive Summary

This deliverable provides an overview of the current urban greening and relevant data ecosystems across the six Urban ReLeaf pilot cities. Understanding these data ecosystems lays the groundwork to identify current and future opportunities for citizen-based observations to complement official data streams. The deliverable starts by examining the global and European vision in urban greening and sustainability and the critical importance of the open data ecosystem. It describes the current state of cities and the urgent need to cope with climate change and several phenomena, such as the Urban Heat Island (UHI) effect, and the degradation of the quality of air and life in urban environments. Furthermore, we explain the term "urban greening", the willingness of the wider public to contribute to this cause, while stressing the complexity to achieve this goal. Key drivers are global and European initiatives as well as research programmes. In this context, we briefly describe the goals of the most relevant flagships and initiatives of the Group on Earth Observation (GEO).

Furthermore, being inspired by the Open Knowledge Foundation, we introduce the necessity of adopting the concepts of not only Open Data but also of Open Data Ecosystems, and, thus, accelerate the transformation of the existing ecosystem from its linear unsustainable shape to circular, where data assets will not "live only once" but simply will be a piece of the wider "data puzzle space". As part of T3.1, ICCS in collaboration with all the city partners map and synthesize the existing data assets of each Urban ReLeaf pilot city, based on the Open Data Institute's (ODI) Data Landscape Playbook (DLP) methodology. The DLP has four well-defined steps (referred to as 'plays') :(i) define the problem, (ii) map the data ecosystem (iii) assess the existing infrastructure, and (iv) assess the policy, regulation, and ethical context. For the first play, a customised 3-tier methodology was formulated to collect insights from Urban ReLeaf city partners and transform them into high-level scenarios, categorised into four use cases. In the subsequent plays of the DLP, key data asset characteristics were identified (i.e., data owners, data formats and interfaces, data location, update frequency, etc.). Finally, in the fourth play, we recorded information regarding the ethical content and the level of openness of each data source. We present our qualitative and quantitative analyses as well mapping the formal and soft value exchanges, with the aim of identifying citizen-powered observations to complement official data streams.

In the final chapter, we attempted to expand the initially identified data ecosystem and present other open-access datasets that could be explored.

# 1 Introduction

## 1.1 Purpose of this document

This document fulfils the objective of Task 3.1 which include a cataloguing of existing data assets and repositories to establish the current state of the data ecosystem across the Urban ReLeaf pilot cities. Furthermore, we identify opportunities for citizen-science based contributions to complement this existing data ecosystem and support wider goals of **sustaining Citizen Observatories and promoting urban greening**.

The existing data assets (geospatial or non-geospatial) potentially contribute to local-scale urban greening policies and strategies as well as possibly wider EU or global agendas (i.e. European Green Deal and UN SDGs). As such, this document is divided into three main sections, namely **(i)** unpacking the Wider EU and Global Objectives, **(ii)** identifying the data assets of the cities, and **(iii)** exploring indicative open-access data sources that serve the objectives of the UR project.

This first chapter outlines the guidelines and core objectives regarding the safeguarding of urban greening and sustainability, underpinning the European and global goals, such as the **European Green Deal**, **GEO** initiatives (i.e., **EO&CS, GEO Trees** and **GEO-BON)** and the **UN-SDGs** with particular attention to the 11th SDG ("*11.7-provide universal access to safe, inclusive and accessible, green and public spaces, in particular for women and children, older persons and persons with disabilities"*[1]). As such, we aspire to adopt both a top-bottom and a bottom-up approach and explore opportunities where the data ecosystem of Urban ReLeaf could contribute to these objectives and thus reflect them into the upcoming strategic plans and future vision of each pilot city.

In the following chapter we introduce Data Landscape Playbook (DLP) and its methodology, which was developed by the Open Data Institute (ODI), to identify and explore the interconnections of the different data assets. Important characteristics that describe the diverse nature of the data itself (i.e. infrastructures, technologies, ethical aspects and regulations) that have been examined. This methodology to establish the current data ecosystem will support the integration of the data assets in the anticipated Urban ReLeaf digital technologies/tools, which will be developed within Work Package (WP)3.

The next chapter outlines the complementarity between the city-specific data ecosystems and several global-scale open-access data sources and technological tools that could further support WP3 tasks. In certain occasions, where these datasets are provided by official European ministries and agencies, certain specific standards and best practices are set that will guide the design and development of the anticipated Urban ReLeaf digital tools. This will ensure compatibility with the selected standards and potentially their long-term inclusion into authoritative data streams, strengthening the curation, validity, usability of the Citizen Science (CS) data sources.



[1] https://ocm.iccrom.org/sdgs/sdg-11-sustainable-cities-and-communities/sdg-117-provide-access-safe-and-inclusive-green-and

## 1.2 Structure of the document

This document is organized into 8 chapters, whose brief description is summarised below:

- **Chapter 1 – Introduction:** Contains a short introduction of this document as well as an overview of the document, and its relation to the relevant activities within the project.
- **Chapter 2 - Unpacking the wider objectives and the role of data ecosystems:** Outlines the objectives of the project and the main guidelines of the EU (European Green Deal) and global initiatives (i.e. GEO and UN SDGs) in the context of urban greening, and the critical importance of Open Data Ecosystems (ODE).
- **Chapter 3 – Introduction to the data ecosystem mapping methodology:** Provides an overview of the methodological steps of the Data Landscape Playbook and the four specific steps (i.e. Plays) that synthesize the ODI-DLP. Additionally, for each *play,* a brief presentation is given.
- **Chapter 4 – Identify the necessities and transform them into scenarios:** Summarizes the 3-tier investigation approach, which tries to uncover the problem and strengthen our connection to the existing data sources, by reviewing the objectives described in the GA and the workshops organised in the contexts of WP2 and WP3.
- **Chapter 5 – Understand the concepts of data ecosystem mapping and assessing the data infrastructure:** Defines the actors that own the different data assets for each city and provides insights on how these data sources can be accessible.
- **Chapter 6 – Access the policy, regulatory and ethical context:** Provides insights into the most essential part of the data spectrum, giving a thorough understanding of the ethical content behind each data source.
- **Chapter 7 – Contribution of the open-access data sources:** Identifies the existing data sources that are available in open data access repositories and initiatives, which can complement the Urban ReLeaf technologies and tools.
- **Chapter 8 – Conclusion:** Summarises the core elements of the document, including some final remarks on its contribution to the related tasks of the project.
- **Appendix A:** Presents visual representation of the Urban ReLeaf city -specific data ecosystem maps.
- **Appendix B:** Presents a summary table of the data ecosystems

# 2 Unpacking the wider objectives and the role of data ecosystems

## 2.1 Global and EU vision in urban greening sustainability

Changes in land cover due to intense urbanisation and spiralling growth of the global population have activated significant change to the local surface balance and atmosphere, leading to the Urban Heat Island (UHI) effect[2]. In addition, financial insecurity, social inequality and in general the deprivation of citizens' living standards[3], further depleted the urban ecosystems; highlighting the importance of considering vulnerable societal groups, such as children, elderly people, people living in poverty and with people with fragile health, in order to design impactful adaptation plans to cope with climate change. The International Panel on Climate Change (IPCC) firmly advocated that nature-based solutions, such as urban greening is an impactful mitigation measure to combat global warming of 1.5°C and climate change from a bottom-up perspective[4]. Such mitigation and adaptation strategies are expected to impact several Sustainable Development Goals (SDGs), such as no poverty (SDG1), good health and wellbeing (SDG3), sustainable cities and communities (SDG11) and climate action (SDG13).

In recent years, the term "urban greening" has become prevalent in urban policy and practice. The latest results have shown a long-lasting degradation of the cities in terms of lack of greening, depletion of air quality and quality of life, and in general inequalities in the distribution of "green" in the social sub-groups that are located in the "affluent" or the "less-fortunate" segments of the city. Consequently, it should be acknowledged that addressing urban greening is very challenging, as cities are described as a multi-parametrical ecosystem with various interchangeable dynamic engines that are characterised by different environmental and socioeconomic profiles and needs. Nonetheless, the goal of increasing urban greening, improving quality of life and promoting environmental sustainability are pathways that each city should choose to follow while avoiding formulating strategic plans that are not being inclusive[5]. To cope with these critical aspects, it is strongly advised that the decision-makers should act decisively to (i) achieve net zero circular cities, (ii) build resilient and sustainable cities; and, (iii) foster inclusivity in the cities. Equitable and environmentally sustainable cities can generate resource efficiency, promote ecosystem restoration, hinder biodiversity loss and perhaps also leverage economic growth. Yet, we are experiencing a trend of introducing policies that in certain cases could leave inequalities untreated, as in general, policy-drivers usually tend to adopt "business-as-usual" models, fragile ecosystems and inadequate development patterns, which mostly serve the global demands[6].

To address these challenges frameworks such as the United Nations Sustainable Development Goals (UN-SDGs and Group of Earth Observations (GEO) Work Programme (WP) and European multi-stakeholder initiatives (e.g. European Green Deal), as well as funding programs (i.e., FP7, Horizon 2020 and Europe) were established[7]. GEO has launched several flagships and initiatives to address the critical aspects related to the preservation of biodiversity and natural environments (GEO Biodiversity Observation Network (GEO-BON)),



[2] https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021EF002016
[3] https://www.mdpi.com/2071-1050/15/6/4996
[4] https://iopscience.iop.org/article/10.1088/1748-9326/ab99ff
[5] https://compass.onlinelibrary.wiley.com/doi/epdf/10.1111/gec3.12459
[6] https://wedocs.unep.org/bitstream/handle/20.500.11822/37413/GEOcities.pdf
[7] https://www.theodi.org/article/what-do-we-mean-by-data-access-initiatives

the subsequent monitoring of forest ecosystem biomass and its change due to urbanisation expansion and disturbance of natural sources of green (GEO-TREES), and of course the essential contribution of citizen science (GEO- CITSCI) and Earth Observation (EO) to address those aspects.

Moreover, the EuroGEO initiative was launched in 2019, as the principal segment of GEO in Europe, aiming to proliferate the use of EO data towards the decisions and measures that are taken for the benefit of humankind and the environment, incorporating space-based, airborne, seaborne or land-based observation systems, including observations coming from citizens[8].

The increasing presence of mobile devices, social media platforms and low-cost instruments, offers opportunities, in which citizens can contribute to the work that was so far done by scientific communities and, thus, have a better level of understanding of what influences their lives. Improving access and democratizing data sources is a way to catalyse a deeper and long-lasting citizen engagement, to build trust and capacity across stakeholders, to succeed in fostering more sustainable outcomes[9]. Eventually, the collaborative contribution of all the actors (i.e., citizens, policymakers, and scientific communities, industrial and social actors) that comprise the cities actors' ecosystem is the only way to achieve the SDGs and prevent the subsequent adverse effects of climate change



## 2.2 The importance of a circular open data ecosystem

Cities and the citizens that comprise the "quadruple helix" of a city's profile (i.e., citizens, public authorities, scientific institutions, and private organisations)[10], nowadays and especially after the evolution of Web 2.0, have substantially benefitted from the power of Open Data (OD). They gain the opportunity to have valuable sources, the so-called "digital-weapons", which assist them to better understand their environment and address the urban challenges. In particular, a wider data availability reveals an untapped opportunity for citizens to comment on public sector decisions and, subsequently, improve transparency and stimulate democracy. As data become more accessible, the value of it increases, as a result.

The Open Knowledge Foundation (OKF)[11] highlights that making data open doesn't guarantee its ability to disseminate the information that is hidden in each data asset. From a wider perspective, citizens are not interested in data, as most of the time, they are characterised by a lack of knowledge or an inadequate technical skills to directly exploit the information that is hidden in them. On the contrary, they often get engaged, when intuitive messages and services are built. In this context, the art of storytelling is increasingly adopted as a promising solution, as narratives can play a crucial role in the process of democratisation and sharing of data science. Stories are able to create appealing messages and thus trigger the direct attention of the audience, emotions, prompt understanding and motivation of citizens to be significantly engaged and devoted to a wider cause[12]. By re-contextualizing the data assets and incorporating them into real-life conditions or challenges, explaining these concepts

---

8 https://research-and-innovation.ec.europa.eu/knowledge-publications-tools-and-data/knowledge-centres-and-data-portals/eurogeo/about-eurogeo_en

9 https://blogs.worldbank.org/governance/multi-stakeholder-initiatives-platforms-collective-governance-development

10 https://www.scivil.be/sites/default/files/paragraph/files/2021-10/Scivil%20Data%20Guide%20-%20EN%20-%20oct2021.pdf

11 https://opendatahandbook.org/solutions/en/Open-Data-Ecosystem/

12 https://www.researchgate.net/publication/337213728_Storytelling_for_narrative_approaches_in_citizen_science_towards_a_generalized_model

through the lens of metaphor and under a simplified framework, a greater level of understanding could be achieved by citizens, that don't have the in-depth knowledge of researchers[13].

Further, a phenomenon that is frequent in a city-level scale, is that data is often hidden and resides in the premises of each stakeholder. As a result, the same data source might be created multiple times, following a specific data structure and technological tools, which fit explicitly the purpose of the few (i.e., agency, ministry, or private company). Hence, huge investments are made for the purpose of a unique goal, with the scalability and curation perspectives of the data assets unexploited and in some cases inapplicable. Even within the government, which is the data leader of Governmental Data Assets (GDA), it seems that usually the data sources are "locked up" in department silos, and thus unable to generating real value.

Addressing these barriers, the key term of the Open Data Ecosystem (ODE) has started to gain a wider appreciation, as it emphasises the need to not only provide free and accessible data assets, but rather "a circular, sustainable, demand-driven environment". In this circular framework, data will not "live only once" but it will be a crucial connector to other existing or future data sources. This challenging goal has a prerequisite, which is to connect the "dots" between the different data owners, as currently, the open data ecosystems are linear and exclusive (Figure 1), and thus data is characterised by a lag in value creation. As a result, federated data distribution models are advised to be established as the basis of the circular ODEs. In this framework, multiple, decentralised data access interfaces, stewarded by multiple organisations, could operate separately, but could be interconnected under a common platform. Thus, disseminating not only the data asset itself, but other insights related to the data collection procedure, practices of quality assessment and evaluation, and technological tools and standardisation methods applied, in order to achieve creating accurate, usable and curated data sources.

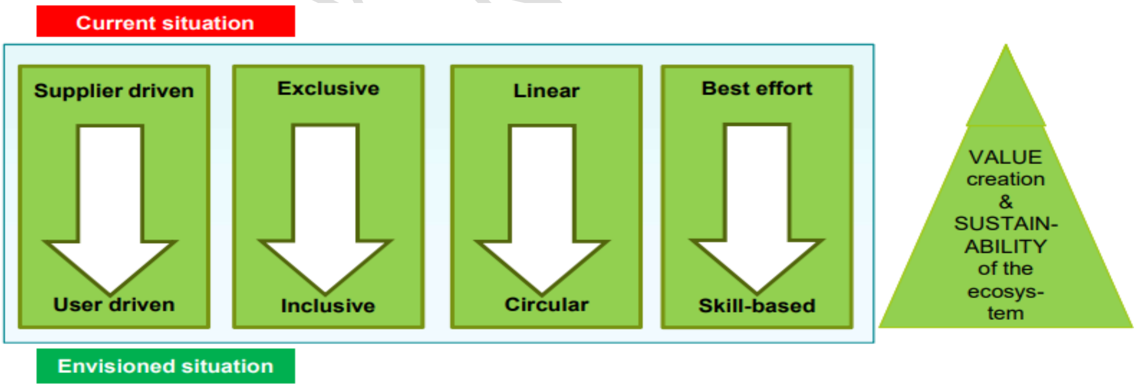



Figure 1: Data ecosystem's current state is described as supplier-driven, exclusive, linear, and best-effort, and the Envisioned situation is characterised as user-driven, inclusive, circular, and skill-based, (source ODECO).

A circular ODE ought to give the opportunity to preserve the data value, so that intermediaries, value-adding participants can keep processing data as long as possible[14]. Nonetheless, in

[13] https://stepchangeproject.eu/once-upon-a-time-the-power-of-storytelling-in-citizen-science/
[14] Alexopoulos, C., Zuiderwijk, A., Charapabidis, Y., Loukis, E., & Janssen, M. (2014). Designing a second generation of open data platforms: Integrating open data and social media. In International Conference on Electronic Government (pp. 230-241). Springer, Berlin, Heidelberg.

order for the ODE to be circular, and able to create a tangible impact on the city, two things are advised to be considered; (i) The necessity to provide a connection with all the above-mentioned stakeholders, who will be promoted to create additional value to the initial content[15], and (ii) the avoidance of a confined circle and thus generating content only in order to cover the needs of limited data providers. Into this circle, both governmental and non-governmental data providers should be integrated. Having the willingness to circumvent this barrier, significant effort should be devoted to stimulate the involvement of all the local problem-solvers in a common and unified ecosystem, in order to achieve a clear view of critical problems that exist, and thus introduce the opportunity for open and broader participation and strategic co-creation.



[15] Charalabidis, Y., Alexopoulos, C., Diamantopoulou, V., & Androutsopoulou, A. (2016). An open data and open services repository for supporting citizen-driven application development for governance. In 2016 49th Hawaii International Conference on System Sciences (HICSS) (pp. 2596-2604). IEEE.

# 3 Introduction to the data ecosystem mapping methodology

## 3.1 Background

This chapter introduces the principles of the Open Data Institute's Data Ecosystem Model (DEM), along with the Data Landscapes Playbook, a well-established, and interactive tool.,. In general, both were designed to support organisations working on "data access initiatives" [16], by providing a set of guidelines to conduct a "data landscape review" and thus understand and derive the most valuable knowledge related to the existing data infrastructure[17], or the data related content, in which they could operate and further expand. The DLP was implemented as part of the Innovate UK-funded R&D project[18] which aimed to help various representatives and data owners (e.g., communities, societies and industries across the world) build their data infrastructure and address common challenges that may appear or have not been foreseen. The DPL was generated through the extensive desk and user research of the ODI, investigating the obstacles and challenges that organizations are still facing when they start formulating a 'data access initiative' and defining their data infrastructure. Key aspects of the DLP include,

- Mapping existing data ecosystems to understand how data is currently being accessed, used and shared, in order to help articulate the vision for the initiative, identify potential areas for additional activities and potential risks.
- Conducting data ethics assessments to help communities work together to identify and manage ethical issues raised by the initiative.
- Creating data inventories identifying the status availability of the relevant data, and specifying information on how data assets are collected, shared, or managed.
- Capitalising or building future data sources on open standards, and thus promoting the collaboration among different organisations within an initiative or across different sectors, to further evolve the initially given data.
- Designing logic models to help organisations to acknowledge the data value they own and facilitate the generation of new opportunities based on the initial inputs and the ultimate impact they can achieve[19].

Leveraging the aspects, in the following chapters, we will attempt to implement the DLP methodology. In fact, we will try to bring the data ecosystem of the pilot cities into light, i.e. (i) identify knowledge regarding the different actors that exist within the city; (ii) describe existing flows among the partners or envisaged flows between the local data holders and the technical partners of Urban ReLeaf; (iii) list the data infrastructures, hosting the data assets; (iv) identify standards, and technologies, that are being used to store, manage and maintain the data assets; (v) describe aspects related to the ethical content using a simplified version of the Data Ethics Canvas (DEC). This deliverable is predominantly oriented on aspects related to the data source, any foreseen limitation with respect to the data curation, and the ethical rights and legislative content around data sources[20].



[16] https://www.theodi.org/article/what-do-we-mean-by-data-access-initiatives
[17] https://theodi.org/project/data-infrastructure-for-common-challenges/
[18] https://open-data-institute.gitbook.io/data-landscape-playbook/master/status-of-this-output
[19] https://www.theodi.org/article/introducing-the-odi-data-landscape-playbook/
[20] https://theodi.github.io/interactive-data-ethics-canvas/

## 3.2 Formulating the guidelines of the Data Landscape Playbook

We start this process by formulating a robust methodological framework that will facilitate the data cataloguing procedure and address objectives related to data heterogeneity, scalability and privacy and data as fit-for-purpose. Towards this direction, we will generate the Data Ecosystem Map of Urban Releaf, capitalising on the **DLP plays**. The DLP consists of t four plays (Table 1) to obtain advice and resources to understand the data landscape and to develop plans to strengthen the data infrastructure, in order to address social, economic and environmental challenges. Under this framework, certain guidelines are necessary to be adopted to apply these steps efficiently.

Table 1: Definition of the Four Players of the Data Landscape Playbook and the steps and tools under which we will attempt to establish this methodology for the Urban ReLeaf needs.

| |
|---|
| Play 1 → **Explore the problem and how data can address it:** provides guidance on how the problem your initiative is seeking to solve can be addressed by designing or strengthening the data infrastructure in order to improve access to data. |
| **Urban ReLeaf approach:** Identify the issues that each city would like to address and how data can help. (Miro-Workshop) |
| **Play 2** → Perform a Data Eco**system Mapping:** Identify stakeholders and map them against the key data and value exchanges within the initiative's data ecosystem. This brings to light the gaps, barriers and opportunities in the data ecosystem, with which your activities can align. |
| **Urban ReLeaf approach:** Use the DEM and create links between the internal data providers and the stakeholders and potential contributors. Discover potential future opportunities. |
| Play 3 → **Assess the existing data infrastructure:** Assess the relevant data infrastructure, including understanding of which data assets you will need, how to access them and how to look for available open standards. |
| **Urban ReLeaf approach:** Collect information from data owners, about data formats, interfaces, and details on data gathering (who, when, how). |
| **Play 4** → **Assess the policy, regulatory, and ethical context:** Understand and be compliant with any ethical context, and assess any legal, regulatory and policy considerations that might impact the data access initiative. |
| **Urban ReLeaf approach:** Guided by the Data Ethics Canvas, provide answers to key aspects, e.g., Data spectrum (Private, Shared, Open), licence, data access copyrights, IPR, legal disclosure statements, or any other legislation, policy or local regulation. |

Through the following chapters, we will try to address the above steps. Starting with the first and most crucial step **(#Play 1)**, we will explore the needs of the respective cities by following a **3-tier investigation approach**. Initiating this process, we will recap the objectives stated in the GA and expected outputs for each city, and how the contribution of the project will be fulfilled. Afterwards, we will extend these objectives with the ones highlighted during the workshop session, (Figure 2) and onwards (i.e., during the interviews performed between the cities and corresponding stakeholders-Task 2.1), with a wider goal to further categorize each city-centred objective and associated assets that are foreseen to be introduced in the Urban ReLeaf lifecycle.

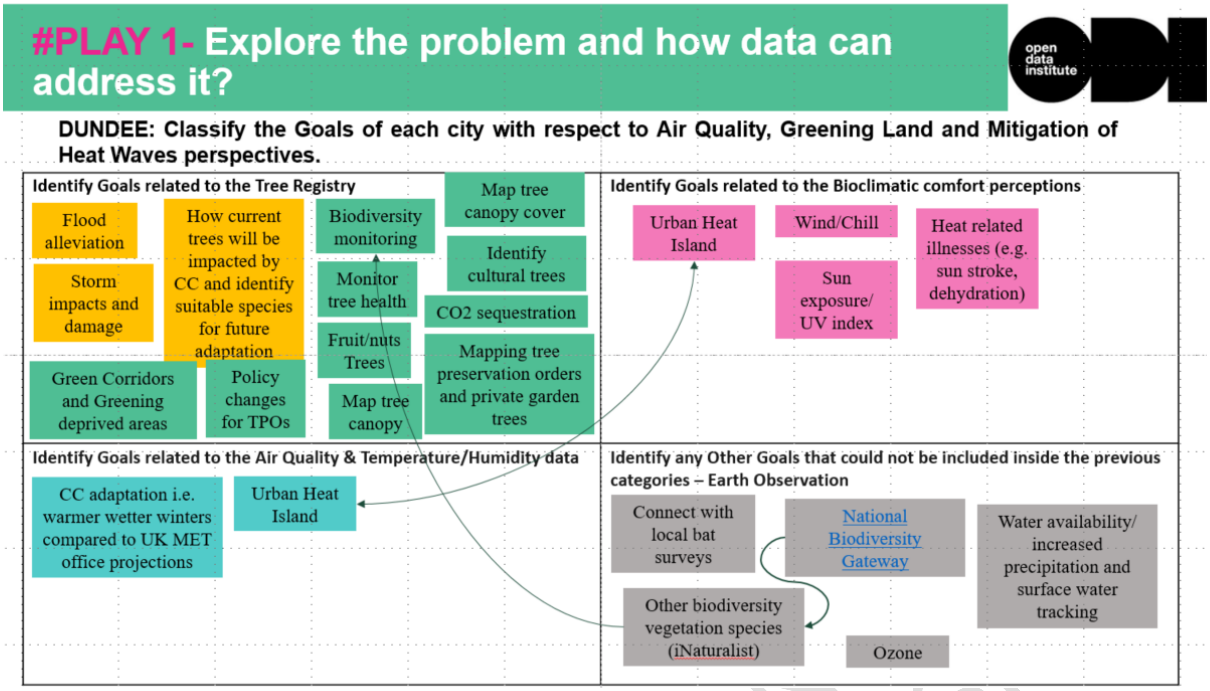



Figure 2: Indicative example of the preliminary goals that have been defined by the city of Dundee, and connections among data sources.

The next play of the DLP requires iterative engagement with multiple stakeholders. Within **Play2: "Data Ecosystem Mapping"**, important information was collected about the general content provided by each data asset, and the actors that are responsible for the ownership of the data. Additional aspects, such as the frequency of the data production and the year of the most recent release, were requested, to identify the availability of the information given and the feasibility of its exploitation within our project. As a result, we will be able to map the main actors of each city, identify flows and exchanges in the ecosystem, and suggest some future insights to be shared between the cities' stakeholders and the consortium partners of this project. An indicative example of the OpenStreetMap (OSM) DEM is presented in Figure 3.

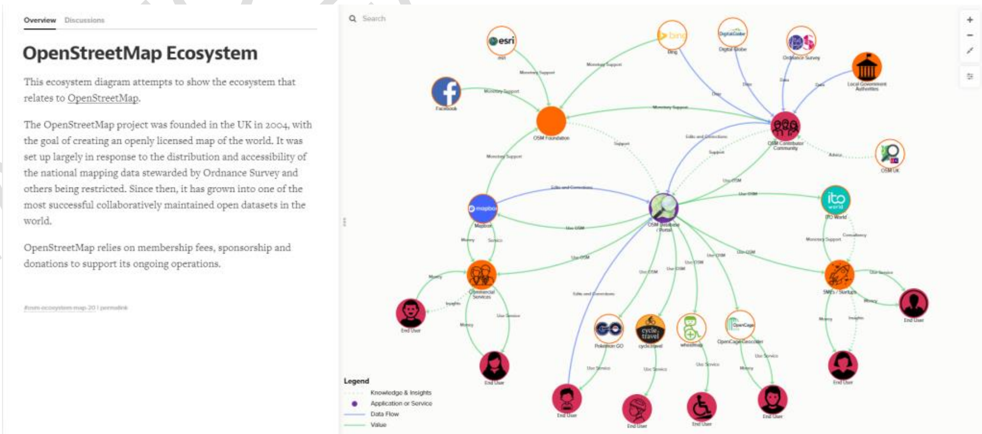



Figure 3: An indicative example of an Ecosystem map. (Image credit: ODI[21] and Kumu[22])

[21] https://www.theodi.org/project/rd-unlocking-the-potential-of-open-geospatial-data-and-technology-in-the-uk/#1539159909718-3299ec7e-3c57
[22] https://kumu.io/ODI/osm-ecosystem-map#osm-ecosystem-map-20

Continuing with **Play3: “Data Infrastructure”**[23], we addressed key aspects related to the data assets themselves, e.g. data formats under which the gathered information is stored (e.g. CSV, GeoJSON, ESRI shapefiles, GeoTIFFs, etc.), standards and technologies that might be used to ensure their curation, data access interfaces (e.g. Web services, APIs, FTPs, Data (Cloud) repositories, etc.).

Finally, within **Play 4: “Assess the policy, regulation and ethical context”**[24], we strived to understand important aspects of the wider impact of data accessing, usage and sharing. The following factors summarise the main aspects that have been requested to be addressed for each data asset.

- **Data spectrum:** Categorising the data asset into the three respective categories of **Closed, Shared, and Open**[25], (Figure 4). In the first case, the data sources are provided upon specified customised licence or by negotiating on a case-by-case basis. Data categorised as “shared” shall be described by pre-emptive licences for specific use-cases, and open data descriptions of both the data and the licence conditions. Finally, when a data source is declared as open, we should verify that is provided under a specific open-accessed licence.

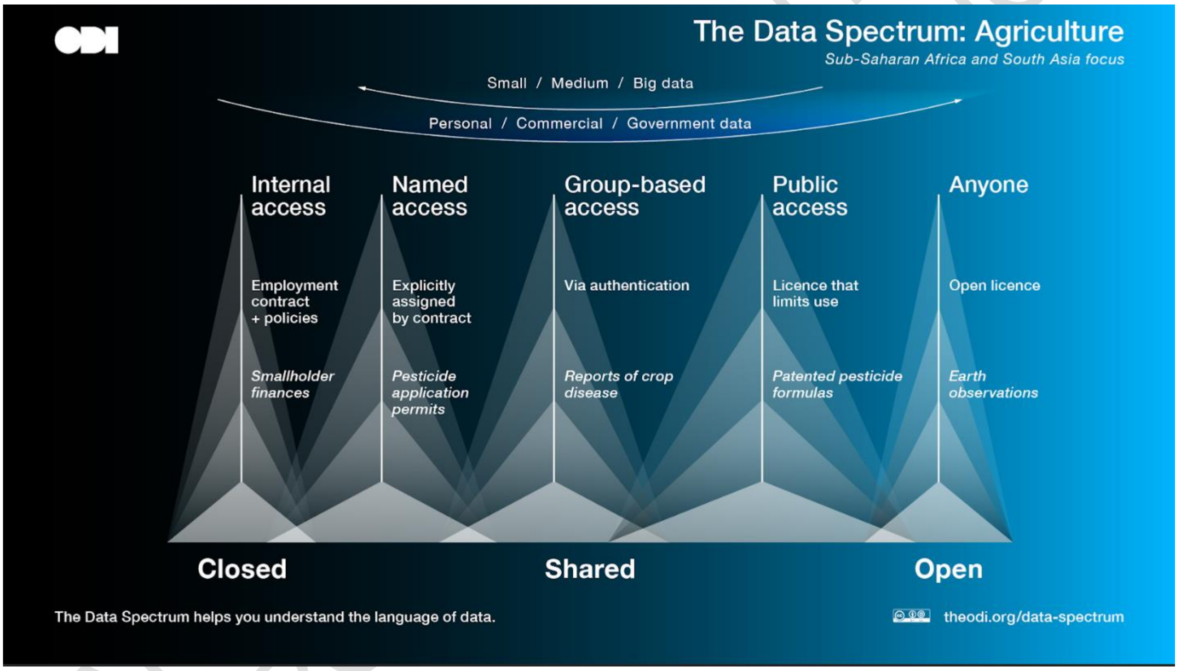



Figure 4: The data spectrum categorisation into Close, Shared or Open, as it has been defined by the ODI, and an indicative example of its application adopting the Icebreaker method in the agricultural sector, (Image credit: ODI, licenced by CC-BY).

- **Limitations:** Consideration of any known bias in data collection, inclusion/exclusion, analysis, or algorithms that strengthen the data quality, gaps that have been ascertained during the data collection process, and therefore any issues that might create a substantial effect on decisions, and thus deviations in the data quality.
- **Ethical and Legislation aspects:** Beyond data protection, it is likely that there are other legislation, ethical frameworks or codes of practice that apply to the sector or

[23] https://open-data-institute.gitbook.io/data-landscape-playbook/play-four-describe-the-data-infrastructure
[24] https://open-data-institute.gitbook.io/data-landscape-playbook/play-three-assess-the-policy-regulatory-and-ethical-context
[25] https://dgen.net/0/2019/05/06/the-data-spectrum-defining-shared-closed/

project. Thus, for each data asset, we specify if there are any laws, data protection guidelines, or data sharing policies, regulations and ethical constraints that might forbid us from accessing the data asset.

- **Openness and Transparency:** Unethical impacts are often the result of a lack of openness and transparency that are in turn a result of unaddressed fears related to data sharing. Thus, in terms of defining the data source as open, specific open-access licences should be declared[26]. Within the Open Data, two popular licences are the Open Data Commons ODbL or Open Database Licence (ODbL) and the Creative Commons (CC0, CC-BY 4, etc.), however, there are other licences that might be formulated in each country[27].

Concluding this analysis, we further support the synthesis of the data landscape, with a literature review aimed to discover datasets that are being provided freely and openly by the research communities, the European and global organisation and agencies (e.g. European Environmental Agency, Forest Information System for Europe, etc.), the EU-funded projects, and CS initiatives (e.g. Geo-Wiki, GROW, etc.). The initial outcomes of this investigation are presented in Chapter 7 of this document and will be further elaborated upon during project implementation.



[26] https://data.europa.eu/elearning/en/module1/#/id/co-01
[27] https://citizens-guide-open-data.github.io/guide/1-open-data

# 4 Play 1- Identify the necessities and the benefits of the contributing data sources

## 4.1 Transforming abstract ideas into scenarios

We help transformed abstract ideas into a series of scenarios for each city that aim to identify needs and make direct connections to anticipated data and technologies. These scenarios help us anticipate future directions and pathways of "how the cities envision the use of the technological solutions that they are offered". At this point, it is important to highlight that in this early stage, such scenarios are an initial basis of ideas and will be refined, through iterative consultations, as the project evolves. This exercise assists both the cities to frame their citizen science campaigns and the technical partners to determine user/technical requirements.

The following sections outline the goals and early draft scenarios of each Urban ReLeaf city. We used a 3-step approach:

- **Step 1:** Recap of the cities' objectives that were defined during the proposal stage (i.e. the GA) and are anticipated to be accomplished.
- **Step 2:** Refine and summarise the goals during early project implementation, which include correspondence during kick-off workshop, monthly 1-on-1 city meetings and linking with outputs from WP2 tasks (T2.1 and T2.2).
- **Step 3:** Transform the above into scenarios, identifying the key actors involved and the envisaged usage of the data and tools.

We intend to address the steps of this methodology with the template presented below (Table 2), which has been employed for breaking down the goals of each pilot into scenarios.

Table 2: Template of the use case and outline of the scenario

| **Use case** | | | |
|---|---|---|---|
| **Scenario** | | | |
| Requested by | | Creation phase | |
| Description | | | |
| Technological Contribution | | | |

The definition of these fields is provided below:

- **Scenario:** Brief outline of scenario
- **Requested by:** The actor who triggered the need of this objective.
- **Creation phase:** Identifies when objective was introduced in the project (Step 1 = proposal phase, Step 2 = early project implementation).
- **Description:** Outlines scenario and workflow based on the early discussions with the city partners.
- **Technological Contribution:** Describes the **external** existing data sources (open-sourced or provided by stakeholder analysis) and technological tools that could facilitate the fulfilment of the goals and contribute to the generation of the expected data (**internal**).

In accordance with this, these goals will be categorized following the defined four use cases:

- **Use case 1:** Participatory Tree Registry (including the Very-High-Resolution (VHR) green layer that will be provided by the satellite Earth Observation (EO) data)**,**
- **Use case 2:** Bioclimatic and subjective perception mapping**,**
- **Use case 3:** Measurement of temperature and humidity, and
- **Use case 4:** Measurement of air quality.

Additional objectives that are of interest to the city that are not incorporated into the four use cases will be represented in a separate group (i.e., Use Case x).

## 4.2 Potential scenarios for Urban ReLeaf pilot cities

### 4.2.1 Dundee

Dundee, one of Scotland's most densely populated cities, declared a climate emergency in 2019. Despite being Scotland's greenest city, the city faces a notable presence of grey infrastructure, and there is an unequal distribution of green and blue spaces in deprived areas. Through Urban ReLeaf, the city of Dundee anticipates **expanding the usability** of existing data sources including the Vulnerability Assessment indicators and the i-Tree canopy inventory of the city to improve the sparse satellite/in-situ datasets and the outcomes of climate change forecasting models. Furthermore, they aim to introduce a 'living' network of observers through citizen-powered science and low-cost sensors (temperatures, humidity, and air quality) to complement official data streams within existing open data portals.

Table 3: Potential scenarios for Dundee

| **Use case 1: Participatory Tree Registry and the VHR green elements monitored by satellite EO data** | | | |
|---|---|---|---|
| **Scenario 1** | **Mapping tree canopy cover (monitor its percentage in terms of covering the national goal)** | | |
| Requested by | DCC | Creation phase | Proposal + Project |
| Description | Existing spatiotemporal coverage of the i-Tree canopy inventory is already available. The i-Tree Eco project has been launched by Treeconomics with the essential contribution of Barcham trees, Forestry of England, and Woodland Trust, and attempts to provide urban forest cover or in general canopy cover of leaves, branches, and stems of trees that cover the ground. Under the urban tree cover GIS platform, users can observe data related to the total canopy cover (%), land area occupied, potential plantable space, historical tree cover and tree cover canopy goals, etc.<br>This information can be accessed and can serve as an essential contribution to the UR tree registry goal, so as the project supports serving the national goal. | | |
| Technological Contribution | Enhance UR tree registry observations; Exploited as a potential feature for the EO-based VHR urban greening elements; Being available through UR data cataloguing service | | |
| **Scenario 2** | **Estimate CO2 sequestration** | | |
| Requested by | DCC | Creation phase | Project |
| Description | Provide estimation of the CO2 annual sequestration (tonnes) per flora taxa through the tree registry app/platform. | | |
| Technological Contribution | UR tree registry, existing inventories and information (e.g. literature review) about the CO2 sequestration per flora taxa. | | |

| **Scenario 3** | **Flood alleviation** | | |
|---|---|---|---|
| Requested by | DCC | Creation phase | Project |
| Description | Stormwater intercepted per year (Gallons) according to the tree type | | |
| Technological Contribution | UR tree registry, statistical information could be displayed by borough/neighbourhood or the selected tree | | |
| **Scenario 4** | **Green space quality-Tree health** | | |
| Requested by | DCC | Creation phase | Project |
| Description | Record different characteristics that are related to trees' health (e.g. leaves are brown, or the tree has been demolished, dry, wet, damaged, cut off, impacted from intense wind or storm, etc.) | | |
| Technological Contribution | UR tree registry collecting location of the tree along with the corresponding photo, and labels in case of the above impacts on the tree's health have been observed. | | |
| **Scenario 5** | **Identify cultural trees** | | |
| Requested by | DCC | Creation phase | Project |
| Description | Tree records, associated with a specific historical event, or are characterised as a natural heritage site (NHS). | | |
| Technological Contribution | Based on the data source of the cultural heritage trees, the UR tree registry app can visualise a storytelling and enhance citizens' engagement. | | |
| **Scenario 6** | **Fruit/Nut production** | | |
| Requested by | DCC | Creation phase | Project |
| Description | Mapping the fruit/nut trees that exist over the city | | |
| Technological Contribution | UR tree registry app, where citizens could mention the presence of Fruit/Nut trees. | | |
| **Scenario 7** | **Mapping tree preservation orders (TPOs)** | | |
| Requested by | DCC | Creation phase | Project |
| Description | All types of trees, except for hedges, bushes or shrubs can be characterized as TPO and protected. Unauthorized works to protect trees are forbidden and could lead to prosecution. TPOs are administered by Local Planning Authorities (LPA)[28]. It might be essential for the corresponding local authority of Dundee to have a GIS map with the conservation and TPO areas and the trees that are present in them, or for the user to have the ability to request for a tree to be included in the TPO area. An example can be seen here (URL). | | |
| Technological Contribution | UR tree registry and existing TPO data source | | |
| **Scenario 8** | **Mapping private trees** | | |
| Requested by | DCC | Creation phase | Project |
| Description | Locate private trees records. | | |

[28] https://www.trees.org.uk/Help-Advice/Public/A-brief-guide-to-legislation-for-trees#:~:text=Tree%20Preservation%20Orders%20(TPOs)&text=All%20types%20of%20tree%2C%20but,by%20a%20Tree%20Preservation%20Order

| | | | |
|---|---|---|---|
| Technological Contribution | UR tree registry presented tree records that are located in private spaces. | | |
| **Scenario 9** | **Biodiversity monitoring** | | |
| Requested by | DCC | Creation phase | Project |
| Description | Detect other vegetation types (e.g. lichen, flowers, fungus, etc.). Integrate local bat surveys. | | |
| Technological Contribution | Explore data sources from iNaturalist, National Biodiversity Gateway, and records from the bats' survey, and visualise them through the UR tree registry platform. | | |
| **Use case 3: Measurement of temperature and humidity** | | | |
| **Scenario 1** | **Heat maps** | | |
| Requested by | DCC | Creation phase | Project |
| Description | Combine the existing temperature observations with the ones collected from the CS to enhance the heat maps that are generated. The outcome can contribute to the monitoring of the UHI and the contribution of green and blue features to the prevention of it. | | |
| Technological Contribution | Low-cost temp/humid sensors, enhancing the existing in-situ network | | |
| **Use case 4: Measurement of air quality** | | | |
| **Scenario 1** | **Air quality monitoring** | | |
| Requested by | DCC | Creation phase | Project |
| Description | Collect CS air quality (PM2.5) measurements | | |
| Technological Contribution | Low-cost air quality sensors | | |
| **Scenario 2** | **Wellbeing monitoring** | | |
| Requested by | DCC | Creation phase | Project |
| Description | Urban monitoring using wearable bioclimatic sensors (temperature, and humidity), and smart applications (Bioclimatic comfort and perception monitoring app), and the mobility data of Strava, to provide insights in the quality of the urban spaces (green, impervious, deprived, etc.) and the experiences of citizens. | | |
| Technological Contribution | Statistical models and methods revealing correlations between the temperature/humidity, air quality, citizens' perceptions, and Strava metro mobility data. | | |
| **Use case X** | | | |
| **Scenario 1** | **Vulnerability assessment** | | |
| Requested by | DCC | Creation phase | Proposal + Project |
| Description | Web GIS dashboard where the existing flood vulnerability maps can be visualised. | | |
| Technological Contribution | UR visualisation dashboard, existing Vulnerability Assessment indicators and datasets, existing vulnerability maps. | | |
| **Scenario 2** | **Integrate UR data steams into the open data platform** | | |
| Requested by | DCC | Creation phase | Project |

| | |
|---|---|
| Description | An interactive GIS interactive platform presenting the collected datasets, and any additional that exist in the city (e.g. SDGs Map[29]). This will provide to the cities a tool, where they could examine the current situation and identify future opportunities for greening. Additionally, the generated CS data should be properly transformed to be compliant with FAIR and INSPIRE specifications, and harmonised, in order to be easily integrated in Dundee's open data platform[30]. |
| Technological Contribution | UR visualisation dashboard and the collected CS data by the four digital tools, open-accessed API. |

### 4.2.2 Riga

Riga city has recognised the need of introducing a more integrated approach in order to cope with the increasing urbanisation and the unrestrained climate change impacts. In the UR project, Riga city wishes to create a constant low-cost CS network of temperature/humidity and air quality sensors, to capitalize on an engaging data-driven decision-making platform and have a more thorough view of the quality of the existing green infrastructure based on satellite EO and Copernicus products, as well as citizens' perception and comfort level, and thus to overcome the heterogeneous and fragmented data information.

Table 4: Potential scenarios for Riga

| **Use case 1: Participatory Tree Registry and the VHR green elements monitored by satellite EO data.** | | | |
|---|---|---|---|
| **Scenario 1** | **Tree records and statements about their health condition** | | |
| Requested by | RPR | Creation phase | Project |
| Description | Record different characteristics that are related to trees' health condition (e.g. leaves are brown, or the tree has been demolished, dry, wet, damaged, cut off, impacted from intense wind or storm, etc.). Notifications from citizens when a tree is partially or completely damaged and needs replacement or to be repaired. Achieve radical increase of citizens' awareness and responsibility | | |
| Technological Contribution | UR tree registry app | | |
| **Use case 2: Bioclimatic and subjective perception mapping** | | | |
| **Scenario 1** | **Green space Air quality** | | |
| Requested by | RPR | Creation phase | Proposal + Project |
| Description | Evaluation of green space air quality according to the notifications received by citizens and the data records received from the low-cost sensors. | | |
| Technological Contribution | Low-cost temp/humid and air quality sensors located in pre-determined positions | | |
| **Scenario 2** | **Citizens' perceptions in the vicinity of green territories** | | |
| Requested by | RPR | Creation phase | Proposal + Project |

[29] https://www.sustainabledundee.co.uk/
[30] https://data.dundeecity.gov.uk/

| | | | |
|---|---|---|---|
| Description | Records of citizens' perceptions regarding close to green and grey spaces, expressing in the long run their perceptions after potential improvements in the green spaces. | | |
| Technological Contribution | Bioclimatic comfort and perception app | | |
| **Use case 3: Measurement of temperature and humidity** | | | |
| **Scenario 1** | **Temperature – Humidity CS data** | | |
| Requested by | RPR | Creation phase | Proposal + Project |
| Description | Collect CS in-situ sensors (temperatures, humidity) and receive quantitative information on the air temperature. These observations could be an essential addition in the improvement of heat severity maps. | | |
| Technological Contribution | Low-cost CS temp/humid sensors and data from the existing network. | | |
| **Use case 4: Measurement of air quality** | | | |
| **Scenario 1** | **Air quality monitoring-validation-calibration** | | |
| Requested by | RPR | Creation phase | Proposal + Project |
| Description | Collect air quality observations and receive quantitative information on the air quality and if it is in line with the standards. | | |
| Technological Contribution | Low-cost air quality sensors and existing national network. | | |
| **Use case x** | | | |
| **Scenario 1** | **Biodiversity monitoring** | | |
| Requested by | RPR | Creation phase | Project |
| Description | Integrate and have available in the data-decision visualization portal the dataset of Latvia ornithology (https://dabasdati.lv/en) | | |
| Technological Contribution | UR visualisation dashboard with the accessible information | | |
| **Scenario 2** | **Connection with existing apps/data sources** | | |
| Requested by | RPR | Creation phase | Project |
| Description | Providing the collected observations of UR under easily and open-accessed interfaces to be possible to integrate with existing platforms, (GIS portal-GEORIGA). The provision of such a valuable information will contribute to future decision planning and the improvement of governance in the municipality. | | |
| Technological Contribution | Open-accessed RESTful API | | |

### 4.2.3 Athens

Starting the analysis for the Athens pilot case, it appears that the long-lasting duration of high temperatures, the extended use of alternative low-priced heating resources (e.g. woods and their substitutes), the modern and unsustainable way of living with the intense use of private vehicles, and the lack of green urban spaces produce a domino effect to the health state of the urban environment. A subsequent degradation is observed with the increase in intensity and duration of heatwaves, the worsening of air quality and health impacts to elderly people, the absence of an updated mitigation plan to cope with climate change related phenomena, such as flash floods and wildfires, as well as the UHI effect of thermal discomfort and heat-related illness (e.g., sunstroke and dehydration).

Thus, to cope with these needs, we anticipate focusing on the following; (i) a comprehensive Athens tree inventory based on existing city-level tree records, observing its association with the air quality, temperature, and humidity microclimate conditions, (ii) an up-to-date and comprehensive data platform in street-level detail for local authorities, including a tool to instantiate tree-related tasks and improve transparency between the different actors and capacity building, (iii) the visualisation of the up-to-date VHR urban greening layer and the contribution to Copernicus Urban Atlas service, and (iv) the expansion of the existing PANACEA network with the incorporation of low-cost, low-footprint miniaturised portable temperature-humidity and air pollution monitoring devices.

Table 5: Potential scenarios for Athens

| Use case 1: Participatory Tree Registry | | | |
|---|---|---|---|
| **Scenario 1** | **Registry with tree records** | | |
| Requested by | DAEM | Creation phase | Proposal + Project |
| Description | Create a complete tree inventory with tree records arising from multiple data sources, such as (i) the outdated existing tree inventory, (ii) the Adopt-a-Tree initiative, (iii) national garden trees registry, (iv) dataset with tree obstacles, (vi) the ones collected within UR project, etc. | | |
| Technological Contribution | UR tree registry mobile app/platform | | |
| **Scenario 2** | **Connecting the green spaces** | | |
| Requested by | DAEM | Creation phase | Project |
| Description | Create a high quality and resolution layer with the location where green elements exist in the municipality of Athens and promote the continuation of greening in the largest roads in order to generate green corridors. | | |
| Technological Contribution | UR tree registry records, VHR EO-driven urban green layers, UR visualisation dashboard (or RESTful API). | | |
| **Scenario 3** | **Digitalisation and Green Management** | | |
| Requested by | DAEM | Creation phase | Project |
| Description | Create a high-quality and resolution green layer with locations of trees and identify areas or sub-regions that lack vegetation, prioritising the municipality's greening initiatives. Provide a digitalised platform, where all this information will be visualized. | | |
| Technological Contribution | UR tree registry records, VHR EO-driven urban green layers, UR visualisation dashboard (or RESTful API). | | |
| **Scenario 4** | **Mapping the private trees** | | |
| Requested by | DAEM | Creation phase | Project |
| Description | Locate private trees records. | | |
| Technological Contribution | UR tree registry presented tree records that are located in private spaces. | | |

| | | | |
|---|---|---|---|
| **Scenario 5** | **Mobile application for tree registry** | | |
| Requested by | DAEM | Creation phase | Proposal + Project |
| Description | Create a usable mobile tool for the collection of the tree records, and a platform with the CS observations. | | |
| Technological Contribution | UR tree registry mobile app and platform | | |
| **Scenario 6** | **VHR green layer** | | |
| Requested by | DAEM | Creation phase | Proposal |
| Description | Capitalise on very high resolution (VHR) multispectral satellite observations and the technical recommendations and specifications for the European Environmental Agency (EEA) and provide a VHR layer (<1m) with the green urban spaces over the municipality of Athens. | | |
| Technological Contribution | VHR EO MSI data, tree records from the tree registry app, open-accessed Land Use/Land Cover (LULC) datasets | | |
| **Use case 3: Measurement of temperature and humidity** | | | |
| **Scenario 1** | **Temperature-Humidity network** | | |
| Requested by | DAEM | Creation phase | Proposal + Project |
| Description | Expanding the existing network with the temperature-humidity sensors. Such an outcome will provide a substantial assistance as it can be further used by future planning practices focused on heat stress reduction. | | |
| Technological Contribution | Low-cost portable temp/humid sensors | | |
| **Use case 4:** Measurement of air quality | | | |
| **Scenario 1** | **High-Resolution BC and UFP exposure maps** | | |
| Requested by | DAEM | Creation phase | Proposal + Project |
| Description | Provide exposure maps at the highest quality and spatial resolution, giving estimations of the Black Carbon (BC) and Ultrafine Particles (UFP) | | |
| Technological Contribution | (i) Miniaturized portable BC (e.g., Aethlabs MA200 micro-aethalometers) and UFP sensors, (ii) Atmospheric composition datasets from the PANACEA and National Air Pollution Monitoring Network (NAPMN)[31] | | |
| **Scenario 2** | **$PM_{2.5}$ in-situ observations** | | |
| Requested by | DAEM | Creation phase | Proposal + Project |
| Description | Unified $PM_{2.5}$ stations and sensor network, including the existing network, such as PANACEA real-time PM2.5 visualisation platform (air-quality.gr) and existing PM2.5 datasets from EU-funded projects, e.g. HackAir[32] and Compair[33]. | | |
| Technological Contribution | UR PM2.5 low-cost sensors | | |
| **Scenario 3** | **Standard calibration methods** | | |
| Requested by | DAEM | Creation phase | Project |

[31] https://emission-project.eu/index.php/en/objective
[32] https://www.hackair.eu/other-initiatives/
[33] https://www.compair.com/en/

| | | | |
|---|---|---|---|
| Description | Perform statistical analysis with the existing in-situ network and the UR low-cost sensors, and laboratory methods, in order to safeguard the provision of curate CS data streams. | | |
| Technological Contribution | Existing in-situ network (e.g. Thissio supersite of NOA) | | |
| **Use case X** | | | |
| **Scenario 4** | **Association between urban green spaces and air quality** | | |
| Requested by | DAEM | Creation phase | Project |
| Description | A common web interface presenting tree registry observations, VHR EO-driven green layers and the temperature-humidity and air quality data streams. | | |
| Technological Contribution | UR visualisation dashboard with the accessible information | | |

### 4.2.4 Mannheim

Mannheim has been described as one of the hottest cities in Germany, with the future climate projections expecting a temperature increase of 2ºC by 2050 and 3-4ºC by 2100, and the intensification of weather events. Anticipating to achieve the climate neutrality 2030 goals and addressing the European Green Deal objectives through the Alliance Local Green Deals, ALLIANCE | Home (localgreendeals.eu), the Mannheim city envisages in the UR project to update their existing but rather obsolete (2014) tree inventory, collecting as well information about species, physical attributes (e.g. texture, form, size, and colour)[34] and locations from both public and private urban lands. Furthermore, under a common and homogenised dashboard, diverse data sources from satellite missions and in-situ sensor networks should be illustrated to assist local authorities in the climate change mitigation planning. Towards this vision, existing local networks, data repositories and projects, such as the SMART City GmbH and SMArt roots projects, and the GIS data platform of Mannheim including several data sources (i.e., replanting trees datasets, SDGs map, cool places, etc.) should be used and further expanded with the contribution of UR outcomes.

Table 6: Potential scenarios for Mannheim

| **Use case 1: Participatory Tree Registry** | | | |
|---|---|---|---|
| **Scenario 1** | **Tree registry app/platform** | | |
| Requested by | MANN | Creation phase | Project |
| Description | Update of the outdated tree inventory (2014) | | |
| Technological Contribution | UR tree registry, existing tree inventory (https://www.gis-mannheim.de/) | | |
| **Scenario 2** | **Mapping private trees** | | |
| Requested by | MANN | Creation phase | Project |
| Description | Locate private trees records. | | |



34 https://edis.ifas.ufl.edu/publication/EP433

| | | | |
|---|---|---|---|
| Technological Contribution | UR tree registry presenting tree records that are located in private spaces. | | |
| **Scenario 3** | **Identification of greening urban gaps and hosting 50,000 trees** | | |
| Requested by | MANN | Creation phase | Project |
| Description | Finding public urban spaces that lack green in order for the city to start planning their planting activities. | | |
| Technological Contribution | UR visualisation dashboard with the VHR urban green layers and collected trees will be presented. | | |
| **Scenario 4** | **Green space quality-Tree health** | | |
| Requested by | MANN | Creation phase | Project |
| Description | Record different characteristics that are related to trees' health (e.g. leaves are brown, or the tree has been demolished, dry, wet, damaged, cut off, impacted from intense wind or storm, etc.) | | |
| Technological Contribution | UR tree registry collecting location of the tree along with the corresponding photo, and labels in case of the above impacts on the tree's health have been observed. | | |
| **Use case 3: Measurement of temperature and humidity** | | | |
| **Scenario 1** | **TRH monitoring and validation** | | |
| Requested by | MANN | Creation phase | Proposal + Project |
| Description | Collect CS in-situ observations of temperature and humidity expanding the existing network of the sMart roots project. | | |
| Technological Contribution | Low-cost Temperature and Humidity sensors | | |
| **Scenario 2** | **Unified sensors network** | | |
| Requested by | MANN | Creation phase | Project |
| Description | Provide low-cost temperature-humidity sensing devices that will be calibrated and validated in order to be easily integrated with the existing network. | | |
| Technological Contribution | Low-cost Temperature/Humidity sensors Calibration/Validation. | | |
| **Use case X** | | | |
| **Scenario 1** | **Integrate existing data sources** | | |
| Requested by | MANN | Creation phase | Project |
| Description | A common web interface that will contain all data sources from the tree registry observations, the VHR EO-driven green layers, and the temperature-humidity sensors. | | |
| Technological Contribution | UR visualisation dashboard with UR generated information and the data from the open-data portal, and the Geoportal of the city of Mannheim, (e.g. city map, noise mapping, solar cadastre, cool places, replanting trees, sister cities, heavy rain hazard map, SDG map, cultural atlas, etc.). | | |

### 4.2.5 Cascais

Important elements for Cascais are to gain a more comprehensive understanding of the impacts of green urban infrastructure in addressing climate change-related phenomena, such as heat waves, extreme flash floods events, biodiversity degradation and the general impact on quality of life. In the UR project, we will capitalise on existing data sources and tools of Cascais, such as six certified meteorological sensors and 9 micro-sensors placed in urban areas, and geospatial data sources from the GEOCascais portal, as well as the bioclimatic comfort and perception tool and the low-cost temperature/humidity sensors built within the UR project, in order to formulate a standard and scalable methodology for measuring the comfort level of the citizens over the UHI effect. These expected outcomes will provide assistance to the city administration to understand how the UHI effect can be tackled in cities' planning, and thus rejuvenate citizens' engagement and build stronger awareness.

Table 7: Potential scenarios for Cascais

| **Use case 2: Bioclimatic and subjective perception mapping** | | | |
|---|---|---|---|
| **Scenario 1** | **Measuring comfort level over UHI** | | |
| Requested by | EMAC | Creation phase | Proposal |
| Description | Standard and scalable tool or methodology for measuring the comfort level of the citizens over the UHI effect and in different urban territories in order to identify differences in the comfort level in impervious and green urban spaces. | | |
| Technological Contribution | UR bioclimatic comfort and perception tool | | |
| **Use case 3: Measurement of temperature and humidity** | | | |
| **Scenario 1** | **Temperature – Humidity CS data** | | |
| Requested by | EMAC | Creation phase | Proposal + Project |
| Description | Collect CS in-situ sensors (temperatures, humidity), and generate an "alive" network of moving observers to the city | | |
| Technological Contribution | Low-cost CS temp/humid sensors | | |
| **Use case X** | | | |
| **Scenario 1** | **Integrate relevant data from Cascais apps** | | |
| Requested by | EMAC | Creation phase | Project |
| Description | GeoCascais is a geospatial repository for storing data from various domains (e.g. Vegetation, Geomorphology, etc.), DataCascais, and DataHub that holds various data sources that are relevant to the project (e.g. Plant trees, species, interventions to trees cultivation, etc.). | | |
| Technological Contribution | Open-access API or the Provision of a common platform. | | |

### 4.2.6 Utrecht

Utrecht city faces a continuous population increase with 400,000 residents to be foreseen by 2028, an outcome which in combination with the intense temperature increase will stimulate a depletion of the quality of life of the locals. Thus, through the UR project, Utrecht city anticipates improving the urban heat stress assessments and eliminating the existing spatiotemporal gaps with the contribution of low-cost CS observations of temperature and humidity. These observations are foreseen to be collected in urban spaces with different characteristics (i.e., impervious and green areas) and in private and public green spaces, and therefore become a valuable input into their heat stress models, and the two platforms of Utrecht (i.e., the Data and Knowledge Hub Healthy Urban Living open access data platform and the Digital Twin of Utrecht). Finally, additional and more personalised insights into the citizens' comfort levels are willing to be covered, as well as an improved mapping of the green spaces based on VHR EO data.

Table 8: Potential scenarios for Utrecht

| | | | |
|---|---|---|---|
| **Use case 2: Bioclimatic and subjective perception mapping** | | | |
| **Scenario 1** | **Bioclimatic comfort monitoring** | | |
| Requested by | PROVUTR; CITYUTR | Creation phase | Project |
| Description | Receive observations on citizens' perception of heat comfort, and its variation over different target groups (e.g. students, elderly people, marginalized groups, and migrant women) and validate the relationship between real-life conditions and the outcomes of the heat stress models over 4 UHI areas. Prioritizing the strategic plan for urban greening based on the measurements. | | |
| Technological Contribution | Bioclimatic comfort and perception app, heat stress models of Utrecht | | |
| **Scenario 2** | **Perceived temperature in private gardens** | | |
| Requested by | PROVUTR; CITYUTR | Creation phase | Project |
| Description | Received citizens' perceptions over heating comfort in private gardens, and showcase the increased heat stress in private properties that lack vegetation and raise awareness. | | |
| Technological Contribution | Bioclimatic comfort and perception app | | |
| **Scenario 3** | **Perceived temperature in heat stress models** | | |
| Requested by | PROVUTR; CITYUTR | Creation phase | Project |
| Description | Integration of the perceived measurements into the heat stress models (e.g. RIVM, KNMI, WUR), or validation of the heat stress models with the perceived measurements. | | |
| Technological Contribution | Bioclimatic comfort and perception app | | |
| **Use case 3: Measurement of temperature and humidity** | | | |
| **Scenario 1** | **Air quality mapping** | | |
| Requested by | PROVUTR; CITYUTR | Creation phase | Proposal + Project |
| Description | Provision of highly accurate CS temperature and humidity data, and generate an alive network of moving observers to the city. Monitoring the effect of blue infrastructures (e.g. ponds) in gardens. | | |
| Technological Contribution | Low-cost CS temp/humid sensors, Snifferbike temperature. | | |

| Scenario 2 | Effect of Green spaces on UHI | | |
|---|---|---|---|
| Requested by | PROVUTR; CITYUTR | Creation phase | Proposal + Project |
| Description | Provide visual evidence of the green features effect (e.g. urban spaces, trees, bushes, and lawns) on UHI areas. | | |
| Technological Contribution | EO-driven VHR green layers, and Low-cost CS temp/humid sensors, showcasing differences in temperature conditions in green and impervious spaces. | | |
| **Use case X** | | | |
| **Scenario** | **Interconnection with digital platforms of Utrecht** | | |
| Requested by | PROVUTR; CITYUTR | Creation phase | Proposal + Project |
| Description | Perform service (e.g. WMS, WFS)/API connections with the existing digital platforms of Utrecht (i.e., Data and Knowledge Hub Healthy Urban Living open access data platform and the Digital Twin) and provide data that will be collected during UR pilot campaigns. | | |
| Technological Contribution | Provide an open-accessed API that could be easily integrated with the Digital Twin of Utrecht. | | |

# 5 Play 2 and 3 – Understand the concepts of Data Ecosystem Mapping and Assessing the Data Infrastructure

## 5.1 Background

Play 1 helped us understand the needs and the potential strategies of the city partners. Having defined these preliminary scenarios, we explored various data sources and open-access portals. In the following chapters, we implement the two following plays of the DLP, which focus on the data ecosystem mapping and the identification the associated data infrastructure. As a result, six different ecosystem maps, one for each Urban ReLeaf pilot city, have been created and presented in Appendix A. For the implementation of these graphical figures, the open-accessed software of draw.io[35] was used.

In the implementation of Play 2 and 3 the following considerations were also addressed:

- **Map the actors:** Identify the local stakeholders of the city (i.e., people, organisations, industries, research institutes, etc.) that are linked to data.
- **Map the "formal" value exchanges:** Draw connections between the actors and the data that they generate. We also expand to depict not only the connections between the actors but, additionally, the data assets and their connections with the stakeholders that are responsible for their stewardship.
- **Identify opportunities and propose "soft" value exchanges:** In order to illustrate the vision of this project, we placed the Urban ReLeaf technical partners in the centre of the DEM and showcase examples of potential connections between the partners and the local data owners. This implies that the technical partner has the ability to collect/enhance/complement that data. For example, the exploitation of heat comfort



[35] https://www.drawio.com/

perceptions and CS temperature observations to enhance the spatiotemporal resolution of the UHI maps.

In the following chapters, we present our findings and describe both qualitative and quantitative analyses. o. During the analysis we categorized each dataset into the four potential use cases. Furthermore, a statistical analysis was performed to examine various characteristics. (Table 9). For this cause, the web-based software Datawrapper[36] was used.

Table 9: Variables for which a quantitative analysis will be performed.

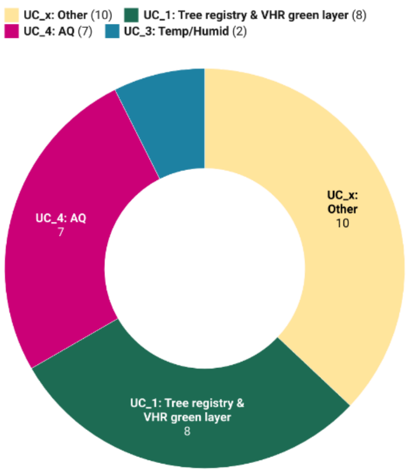
| **Data Ecosystem Mapping related parameters** | |
|---|---|
| **Data assets per use case** | Examine the tendency of the city to address aspects related to the (i) tree records and their conditions, (ii) thermal conditions, (iii) air quality, (iv) climate change, and (v) any other data domain. |
| **Data assets per data owner** | Analyse the volume of the data sources each data owner provides |
| **Data Infrastructure related parameters** | |
| **Data formats** | Examine the common types of data formats, which the data owners usually choose to generate and store their data sources. |
| **Data access interfaces** | Examine the interfaces under which the data sources are maintained and disseminated to the public. |
| **Date of data release** | Provide insights into the tendency of the data sources to be up to date. In terms of data harmonisation, proper aggregations will be implemented and applied to every city's data records. The data records will be classified based on the year of their release. In cases where a data asset is updated annually or within a specific time frame (every 2 or 4 years), multiple records will be kept as an additional record for every year of its creation. This approach is as an efficient solution for the annual updates, and not for the data sources that are updated every hour or within a few days, which would lead to an unprecedented workload, and is out of scope of this deliverable. |
| **Frequency of update** | Tendency of datasets to be continuously updated. |

## 5.2 Discover the Data assets

### 5.2.1 Data Ecosystem Mapping of Dundee

As it has been stated by Greenspace Scotland[37], a good quality green space is a “fit-of-purpose” objective, as it envisions an accessible, safe, welcoming perspective. From our data ecosystem mapping process revealed a vast inventory of assets that address the monitoring the quality of green spaces and air, with their wider goal to safeguard and preserve the natural resources, and as a result further enhance human health, well-being and the QoL for urban residents. (Figure 5).

[36] https://www.datawrapper.de/
[37] https://www.researchgate.net/publication/341678100_A_GIS-based_assessment_of_green_space_accessibility_case_study_of_Dundee




| UCs | Data assets |
|---|---|
| UC_1 | Tree Preservation Order (TPO) Boundaries; National Forest Inventory Woodland Dundee; Tree locations; Survey on Native Woodlands and Plantations on Ancient Woodland Sites of Scotland; Ancient Woodland Inventory; Habitat and Land Cover Map; Local Nature Reserves |
| UC_3 | Ambient and Modelled Temperature |
| UC_4 | AQ NO2, NO, PM[1,2.5,10], SO2; Air Quality Management Area (AQMA); PM2.5 maps; NO2 projections; |
| UC_x | AQ Wind speed and direction; Noise maps; Local Plan Review - Open spaces; National Soil Inventory; Soils Dundee; Scotland Topsoil Organic Carbon Content (TOC); Special Protection Area (SPA) Dundee; Site of Special Scientific Interest (Scotland) Dundee, National Vegetation Classification (NVC) Dundee; RAMSAR Dundee; Scottish Index of Multiple Deprivation |

Figure 5: (Left-side): Number of datasets per Use Case; (Right-side): Brief description of the identified datasets classified in each category.

More specifically, the Dundee City Council (DCC), the Scottish Government, as well as several sub-governmental units (i.e., Scottish Forestry, and Forestry Dundee city council) and statutory corporations (i.e., Scottish Water and Forestry Research) were identified as the main creators of datasets. The datasets themselves focused on the green level of the city, including details on public tree locations, and additional information of the flora species, the extensive forest inventory in woodlands and plantations, with a significant interest in trees of cultural value. Furthermore, it appears that the DCC has supported the subsequent identification of trees that make a special contribution to the landscape of the city. , The TPO geospatial records includes trees or group of trees that have been declared to have particular amenity value and with a necessity to be under special protection by the local authority[38]. In a similar view, the city generates a geospatial local-scale layer with the areas that are characterised by a natural heritage importance. In a wider perspective, habitat and land cover maps have been created with the contribution of Artificial Intelligence (AI) models and spaceborne remote sensing observations.

The National Department for Environment Food and Rural Affairs (DEFRA) and the Scottish Environment Protection Agency (SEPA) are the main actors responsible for the monitoring of AQ , measuring various parameters, such as wind speed and modelled wind speed, wind direction and the corresponding simulated results, Nitrogen dioxide and oxide (NO2, NO), Modelled and Ambient Temperature, Particulate Matter (PM) 1, 2.5 and 10, Sulphur dioxide (SO2), and decades of others at over 1500 sites across the whole United Kingdom (UK). Complementary to these, NO2 and PM2.5 future estimations, are general viewpoints on the Air Quality Monitoring Area (AQMA) provided by the Scottish Water organisation, expressed in rasterised values. Additionally, the Scottish government provides on a five-yearly cycle, noise geospatial maps to help assess quality of life in Dundee.  This innovative work will include new road and rail noise source data, and a new calculation methodology (CNOSSOS) to significantly improve the evidence base and follow the guidelines of the Environmental Noise Directive 2002/49/EC[39].



38 https://www.glasgow.gov.uk/index.aspx?articleid=26026#:~:text=The%20TPOs%20are%20made%20to,potential%20threat%20to%20the%20trees.

39 https://noise.environment.gov.scot/

Eventually, other geospatial data layers are generated to better monitor urban sustainability and biodiversity or further improve them in the future. In particular, the James Hutton research institute is devoted to the investigation of soil health and thus provides essential information over several parameters that monitor its health conditions, with an indicative example being the Topsoil Organic Carbon Content (TOC), not explicitly for the Dundee city but for the greater area of Scotland. Furthermore, several geospatial layers are produced that aim to depict the adequacy of open green spaces in Dundee for potential renewal. Also, the level of deprivation, the sites of special scientific interest, and national vegetation classification, and the areas under a Special Protection Area (SPA), expressed by two layers; (i) SPA and (ii) the Ramsar sites, which are wetlands of international importance designated under the Ramsar Convention[40].

The number and types of data for each data owner are illustrated and described in Figure 6 and Table 23, respectively.

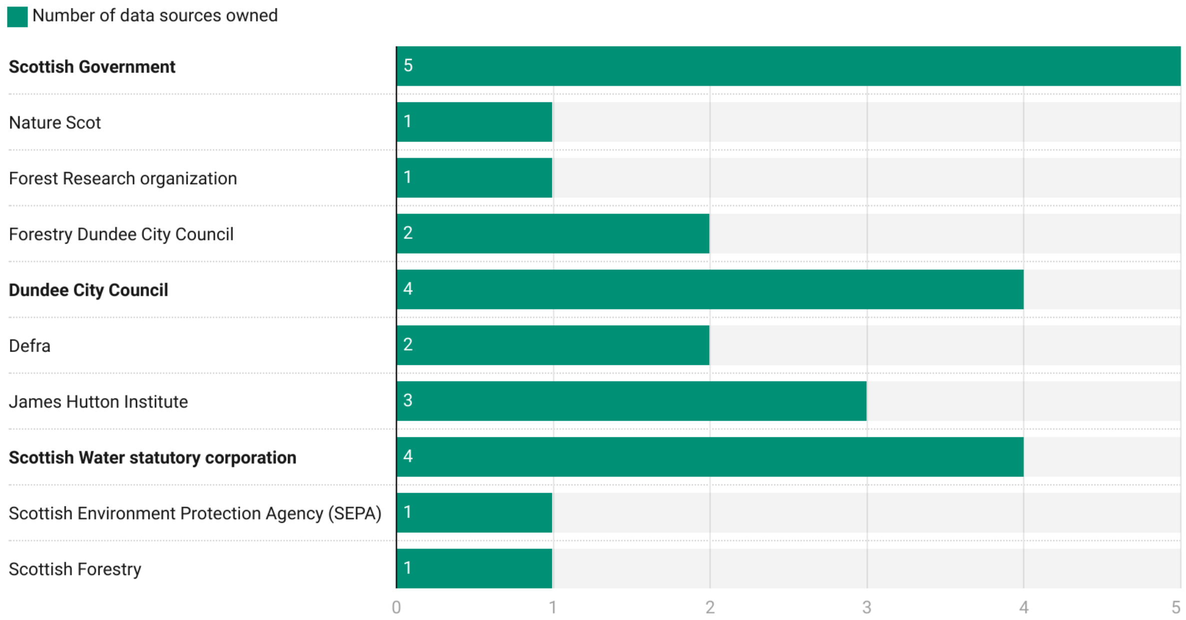



Figure 6: Number of datasets per identified data actor of Dundee.

Table 10: Datasets of the different identified data actors of Dundee

| Data owners | Data assets |
|---|---|
| Scottish Government | AQ: NO2; NO; PM[1,2.5,10]; SO2; Air Quality Management Area (AQMA); Special Protection Area (SPA) Dundee; Site of Special Scientific Interest (Scotland) Dundee; Scottish Index of Multiple Deprivation (SIMD); National Vegetation Classification (NVC) Dundee; Woodland inventory; Scottish Index of Multiple Deprivation |
| Nature Scot | RAMSAR Dundee |
| Forest Research organization | National Forest Inventory Woodland Dundee; Local Nature Reserves; Tree locations; |
| Forestry Dundee City Council | Local Nature Reserves; Tree locations |

40 https://www.ramsar.org/

| | |
|---|---|
| Dundee City Council | Local Plan Review - Open spaces; Tree Preservation Order (TPO) Boundaries; Local Nature Reserves; Tree locations |
| Department for Environment Food and Rural Affairs (DEFRA) | AQ: NO2; NO; PM[1,2.5,10]; SO2; Air Quality Management Area (AQMA) |
| James Hutton Institute | National Soil Inventory of Scotland; Soils Dundee; Scotland Topsoil Organic Carbon Content (TOC) |
| Scottish Water statutory corporation | Habitat and Land Cover Map; Air Quality; NO2 Projections |
| Scottish Environment Protection Agency (SEPA) | Scotland Noise Map |
| Scottish Forestry | Survey on Native Woodlands and Plantations on Ancient Woodland Sites of Scotland |

Addressing the data formats (Figure 7) , geospatial information is collected by the data owners, in which the most known options such as (i) the popular JavaScript-based data format (JSON) and its derivative (GeoJSON) that allows the effortless integration with web-based applications, (ii) the Keyhole Markup Language (KML) is an XML-based data format introduced by Google in Google Maps and Google Earth, the (iii) ESRI shapefile explicitly for the vectorised data sources, which is the native format for the ArcGIS software suite from ESRI, which has been adopted by every subsequent major GIS systems, and the GeoTIFF format for the rasterised products. Rather than that, we could also mention that the majority of the data owners are operating using the ArcGIS enterprise suite, as there are additional data formats, e.g. Geospatial Database (GDB) file and the Map Info (.mxd file) that are exclusively produced by the ArcGIS desktop software (SW). Furthermore, the comma-separated values (CSV) and the American Standard Code for Information Interchange (ASCII) file formats were chosen as the most favourable options for meteorological and climatological data observation.

Considering the data storage aspect and the data accessibility (Figure 8), the majority of the identified datasets are disseminated upon open data portals and catalogues, with the Open Geospatial Consortium (OGC) Web Map and Feature (WMS and WFS) services to be offered as an alternative machine-friendly option. When it comes to interface implementation, only one data source is provided through a RESTful API.

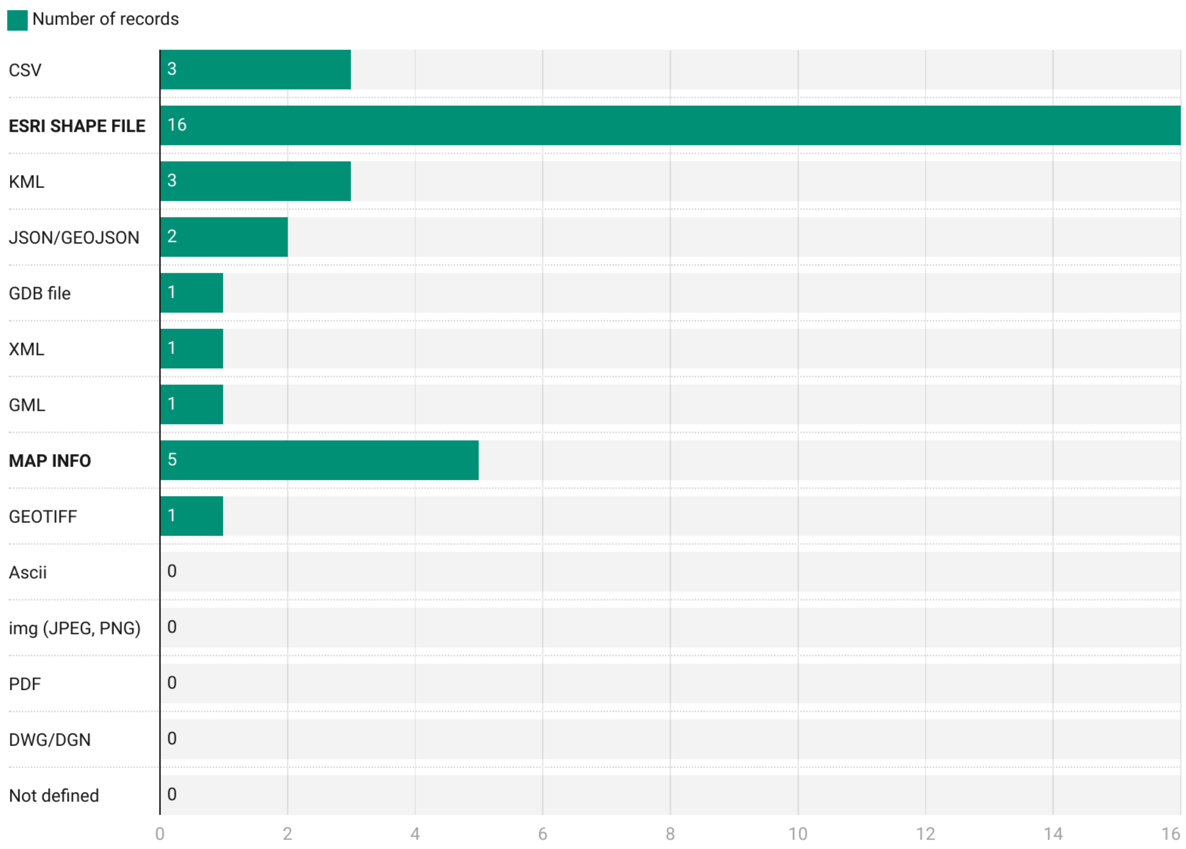



Figure 7: Chosen data format to store the collected data assets.

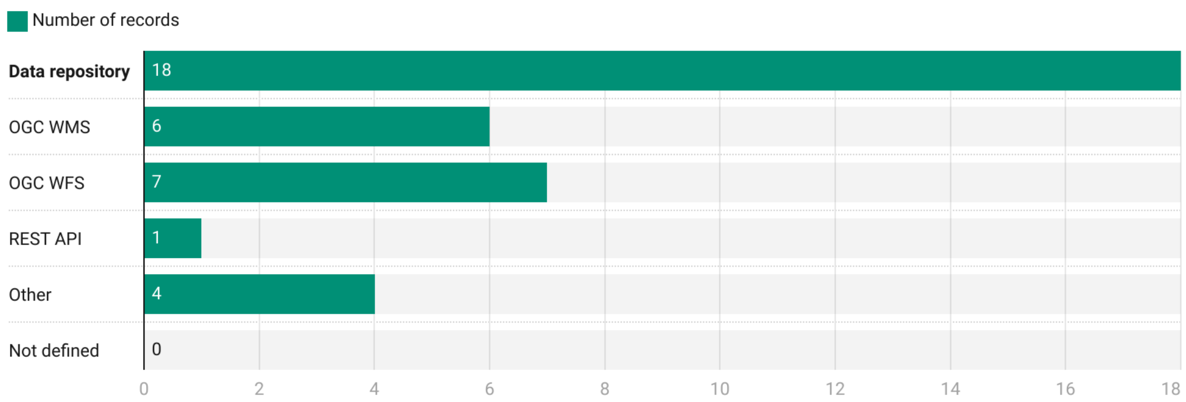



Figure 8: Chosen interfaces (User or Machine-readable), under which the data is available.

The final aspect that we would like to cover refers to the tendency of the organisations to create updated content and actively maintain the data sources. We performed a preliminary analysis on data records that have been published from 1970 until now, using user-determined time intervals. The majority of the data assets have started to be gathered from 2000 and onwards, with the period 2021 and 2022 to stand out. In terms of the intention to maintain such valuable data sources, it doesn't seem to be actively supported as the majority of them either are not defined, or the datasets are updated irregularly.

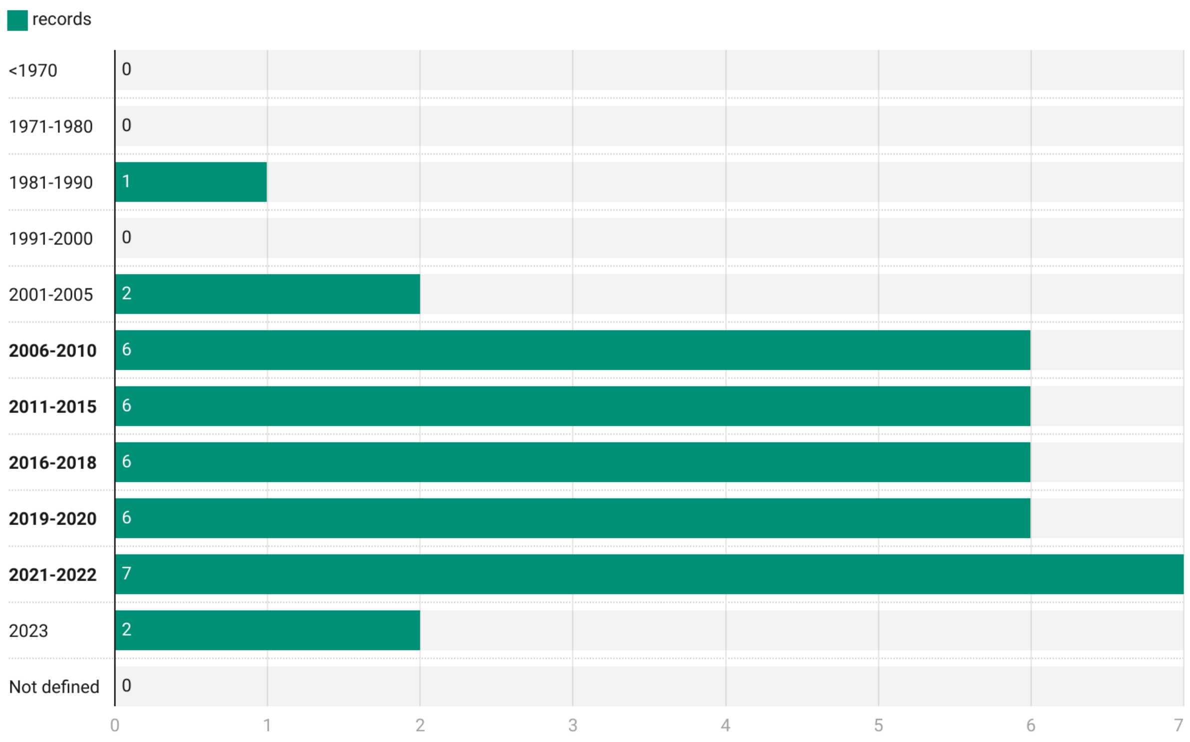



Figure 9: Distribution of the number of data sources per the user-determined time intervals.

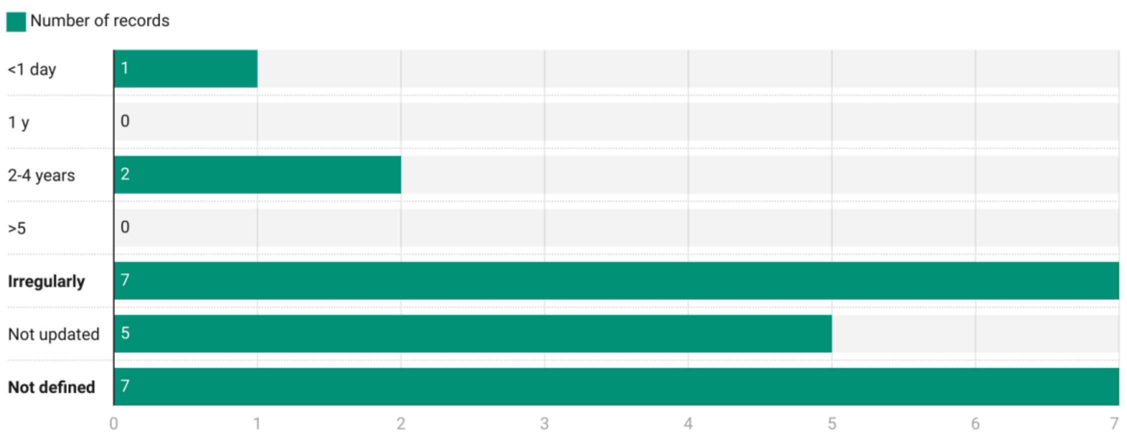



Figure 10: Distribution of the number of data sources per the user-determined classes of update frequency

### 5.2.2 Data Ecosystem Mapping of Riga

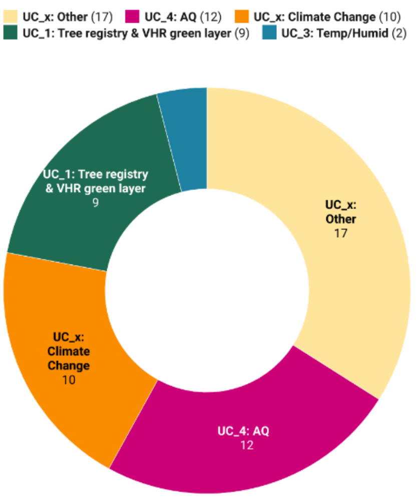
The United Nations Economic Commission for Europe (UNECE) has declared that Riga is undergoing a process of metamorphosis[41], as it has already conceived an innovative strategic plan for a rapid change that allowed to evaluate of the QoL and, at the same time protecting the environment and national heritage, and providing equal opportunity for individual development. With wider ambition, special attention was given to the goals of the European Commission (EC) and the UN-SDGs, and thus a programme was established in 2014[42], which is summarised in the four following strategies[43], listed below. Additionally, Riga has joined the

[41] https://unece.org/fileadmin/DAM/env/europe/workshop/riga.e.pdf
[42] https://www.rdpad.lv/wp-content/uploads/*2014/11/ENG_STRATEGIJA.pdf
[43] https://eurohealthnet-magazine.eu/green-cities-how-riga-is-paving-the-way-towards-healthier-and-greener-urban-areas/

"Green City Accord"[44] initiative of the EC; a movement that has the intention to assemble European mayors that are committed to making cities cleaner and healthier.

- Convenient and environmentally friendly commuting within the city.
- Urban environments promote the quality of life.
- Healthy, socially inclusive, and supportive city to ensure health promotion, improved access to health care, and social support for those at risk of social exclusion.
- Good environmental quality and a resilient urban ecosystem to mitigate climate change.

Targeting such highly impactful goals, Riga has put a tremendous effort to better monitor the quality of several features of an urbanised environment; an outcome that can be easily verified from our data mapping process. Figure 11 depicts the results of the aforementioned procedure, where on the left side the pie chart represents the distribution of the datasets alongside the main four uses of the project and the two complementary ones that include the climate change related datasets and any other data assets that were brought up in the project, and on the right side the table attempts to give an in-depth representation of the exact data sources.




| UCs | Data assets |
|---|---|
| UC_1 | Protected areas, deposits of protected species protected trees; Contaminated sites; Distribution of invasive tree species in Latvia; Sentinel-2 spectral indices; The quality and availability of green spaces in neighbourhoods; Territory Plan of the historical center of Riga (land use and its protection zones); Orthophoto RGB + Infrared Protected trees; Forest parks; Tree registry in red lines, and in educational territories; Cities meadows Dataset from project "City meadows" |
| UC_3 | Air temperature forecasts; Meteorological operational data; Meteorological locations; Meteorological parameters; Meteorological archive data; |
| UC_4 | Air Quality (Index); 5 air monitoring stations and 2 portable measuring PM (dust), Volatile Organic Compounds (Benzene, NO2, and SO2); Zone map for NO2 and PM10 from ADMS Urban 5.0 pollution dispersion modelling; SensorCommunity Data (Luftdaten airRohr PM 2.5); Air quality 3 stations; Monthly AQ reports; AQI |
| UC_x | AQ variables wind direction and speed; Noise pollution; Noise maps; UHI maps; CC projections of 2 GHGs – RCP 4.5 and 8.5 scenarios; Climate portal for Riga municipality (e.g. wind direction, and total precipitation); UV radiation, pollen concentration; Contaminated sites; Surface water quality monitoring stations; Hydrological operational data; stations; parameters [Water level, Water temperature, depth]; archive data; Statistical records; Population density. |

Figure 11: (Left-side): Number of datasets per Use Case; (Right-side): Brief description of the identified datasets classified in each category. The data not provided with a hyperlink is either closed or the city didn't provide any and we couldn't identify any after our research on the Web.

Starting with the data sources that are associated with the tree registry tool, the municipality of Riga along with the two additional departments of Building Control and the Housing and Environment of Riga city council are the main contributors to the following data assets: geospatial records of trees across the red lines and in education territories, forest parks, as

[44] https://environment.ec.europa.eu/topics/urban-environment/green-city-accord_en

well as trees and areas of special importance, the quality and availability of green spaces in neighbourhood, meadow locations gathering under the framework of Riga City Council's project "City meadow"[45]. Additionally, examining factors such as EO-based vegetation indices contribute to datasets that showcase the quality and availability of green spaces in the neighbourhood. The geospatial plan of the historical centre of Riga as an outcome of the general renewable process and sustainable development contains information related to the distribution of the land use types and the protection zones. Under a similar perspective, the Nature protection administration, subordinate to the Ministry of the Environment Protection and Regional Development is responsible for the provision of datasets that present the protected areas, where both rare flora and fauna species exist and the distribution of invasive tree species.

Moving towards the air quality and bioclimatic aspects, it seems that noise monitoring has gained significant importance and thus a company in the private sector, namely Latvian State Roads (VSIA), is responsible to collect noise observations and create extrapolated geospatial representations of the noise quality level. Also, both operational and archived meteorological observations (e.g. wind speed and direction, precipitation, humidity and temperature) are provided by the second main data owner of Riga, the Latvian Centre for Environment, Geology and Meteorology. The same contributor along with the Freeport of Riga (Environmental Unit), and the municipality of Riga along with its sub-unit of the department of housing and environment hold more than ten sensors of AQ, which provide certified observations of several air pollutants, such as dust, Volatile Organic Compounds (Benzene, $NO_2$, and $SO_2$) and in locations described with different land use characteristics (e.g. industrial and residential areas at the vicinity to main transportation roads). They are further enhanced by the existing crowdsourced network of sensors' community that has worldwide coverage of over 15k low-cost, do-it-yourself (DIY) PM2.5 sensors. Furthermore, additional geospatial products are produced related to $NO_2$ and PM2.5 zone maps using the advanced Atmospheric Dispersion Modelling System (ADMS) Urban pollution model[46], and therefore, parameters important for human health such as the ultraviolet radiation, the pollination concentration and the AQI indices.

Data sources relevant to climate change are of major concern to Riga, and thus EO-driven variables for monitoring the UHI effect over the city and spatially climate change predicted changes in the climate parameters over time following two climate change scenarios (i.e., Representative Concentration Pathway (RCP) 4.5 and 8.5) are available through a specific climate change analysis tool[47]. Statistical variables for the population density and the gross-salary capacity are about to conclude the overall amount of data assets that exist within this city. With the above review of the identified data assets that comprise the data ecosystem map of Riga to be covered, Figure 12 and Table 11 aim to provide an overview of the distribution of the data assets in each of the identified data owners, (Figure 12), and therefore trace their relation with each specific dataset, (Table 11).



[45] https://eng.lsm.lv/article/society/environment/riga-to-create-20-urban-meadows.a409429/
[46] http://www.cerc.co.uk/environmental-software/ADMS-model.html
[47] https://www4.meteo.lv/klimatariks/

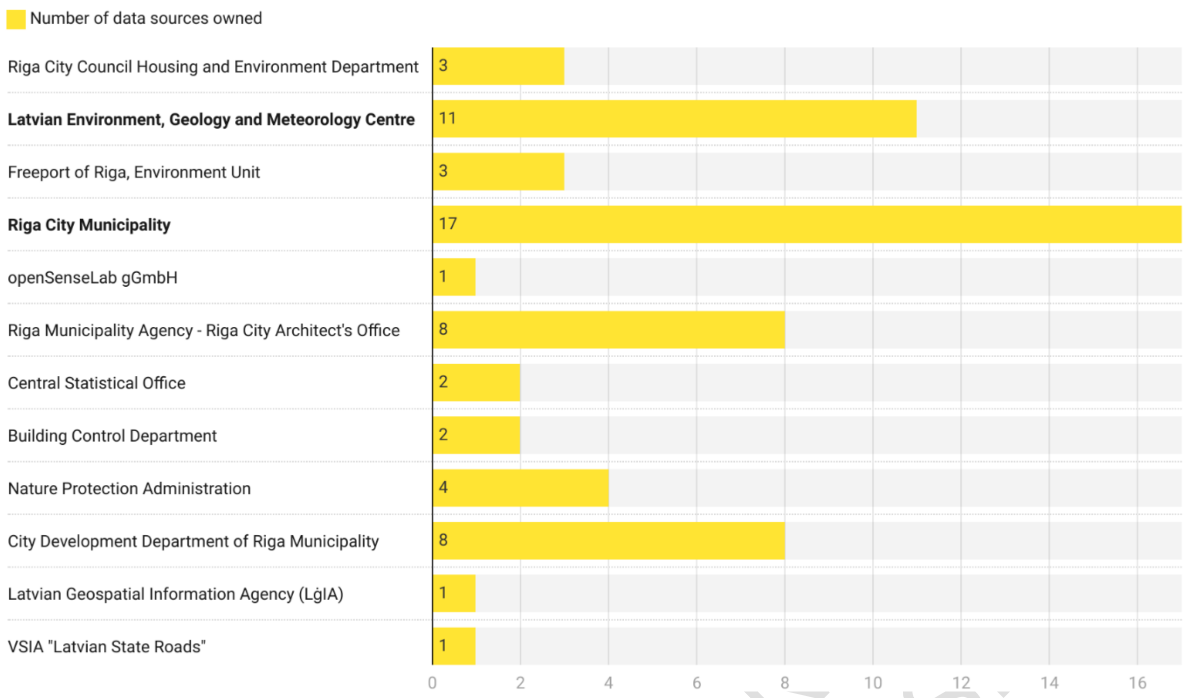



Figure 12: Number of datasets per identified data actor of Riga.

Table 11: Datasets of the different identified data actors of Riga.

| Data owners | Data assets |
|---|---|
| Riga City Council Housing and Environment Department | Forest parks; Tree registry in red lines, and in educational territories; AQI; |
| Latvian Environment, Geology and Meteorology Centre | UHI maps; CC projections of 2 GHGs – RCP 4.5 and 8.5 scenarios; Air quality 3 stations; Contaminated sites; Surface water quality monitoring stations; Hydrological operational data, stations, parameters [Water level, Water temperature, depth]; Meteorological data [operational, locations, parameters, archive]; Climate portal [wind direction, and total precipitation] |
| Freeport of Riga, Environment Unit | 5 air monitoring stations and 2 portable [dust, Volatile Organic Compounds (Benzene, NO2, and SO2]; AQ [wind direction and speed] |
| Riga City Municipality | Noise pollution; Noise maps; Sentinel-2 spectral indices; The quality and availability of green spaces in neighbourhoods; Territory Plan of the historical center of Riga (land use and its protection zones); Cities meadows Dataset; Zone NO2 and PM10 (ADMS); AQI, UV radiation, pollen concentration |
| openSenseLab gGmbH | SensorCommunity Data (Luftdaten airRohr PM 2.5) |
| Riga Municipality Agency - Riga City Architect's Office | UHI maps |
| Central Statistical Office | Statistical records; Population density |
| Building Control Department | Protected trees |
| Nature Protection Administration | Protected areas, deposits of protected species protected trees |
| City Development Department of Riga Municipality | Noise pollution; Noise maps; Sentinel-2 spectral indices; The quality and availability of green spaces in neighbourhoods; |

| | |
|---|---|
| | Territory Plan of the historical centre of Riga (land use and its protection zones); Cities meadows Dataset; |
| Latvian Geospatial Information Agency (LģIA) | Orthophoto RGB + Infrared |
| VSIA "Latvian State Roads" | Noise maps |

Moving towards the additional aspects related to the commonly adopted data formats and interfaces that are of use in Riga city, we could see that CSV is in Riga's data owners' favour; a rather logical result as several AQ/meteorological variables are monitored. With the contribution of the EO-based products, GeoTIFF is applied for the rasterised data sources. Continuing with the more web-based data formats, JSON and its derivative GeoJSON seems to be selected in 6 cases as a complementary method of storing, with the second often accompanying the vectorised ESRI shapefiles. As alternative options, the ArcGIS database file and the map info output products are utilised for the first to be capable of storing multiple data. Lastly, only on a single occasion, the static format of PDF was chosen. 13 datasets require further investigation as the formats are currently unknown.

Considering the data storage interfaces, the majority of the identified datasets are published in open data portals and catalogues, with the OGC-compliant web services being the second favourable option. A single preference appeared on RESTful APIs, whereas for the same 13 records, reference to the interfaces that are utilised was not defined.

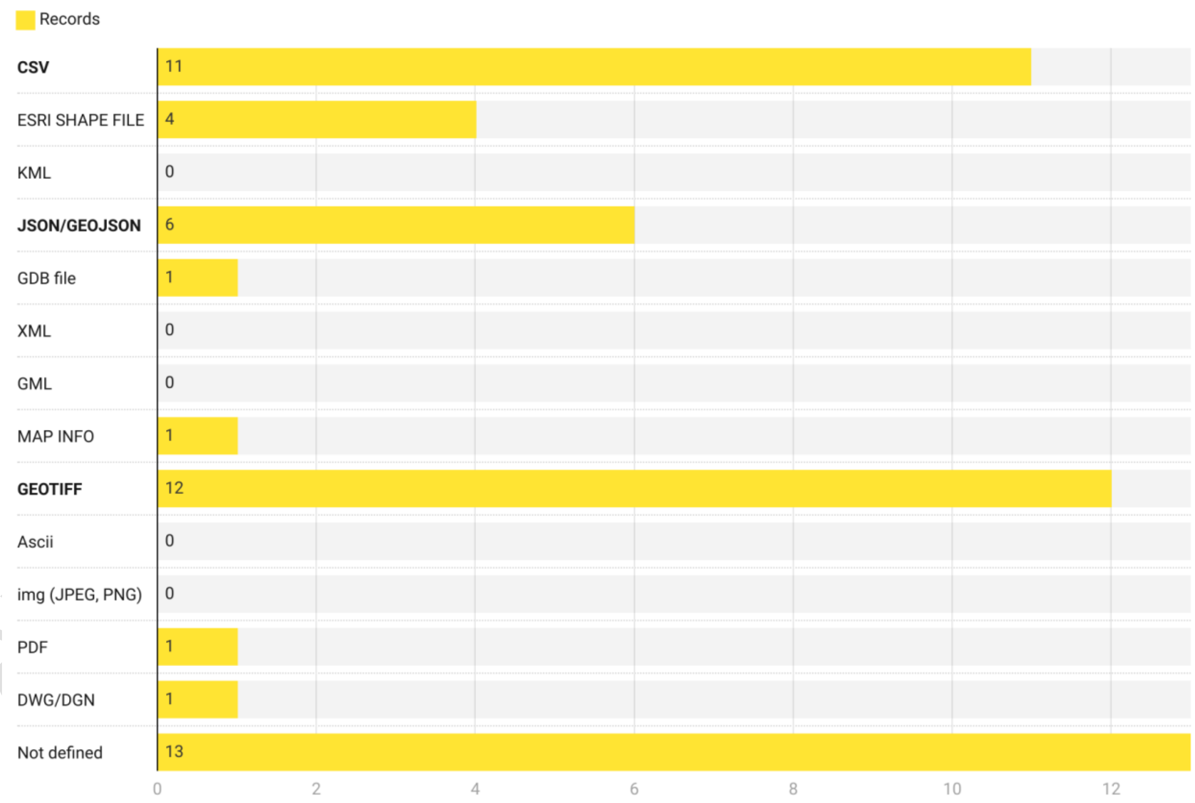



Figure 13: Chosen data format to store the collected data assets.

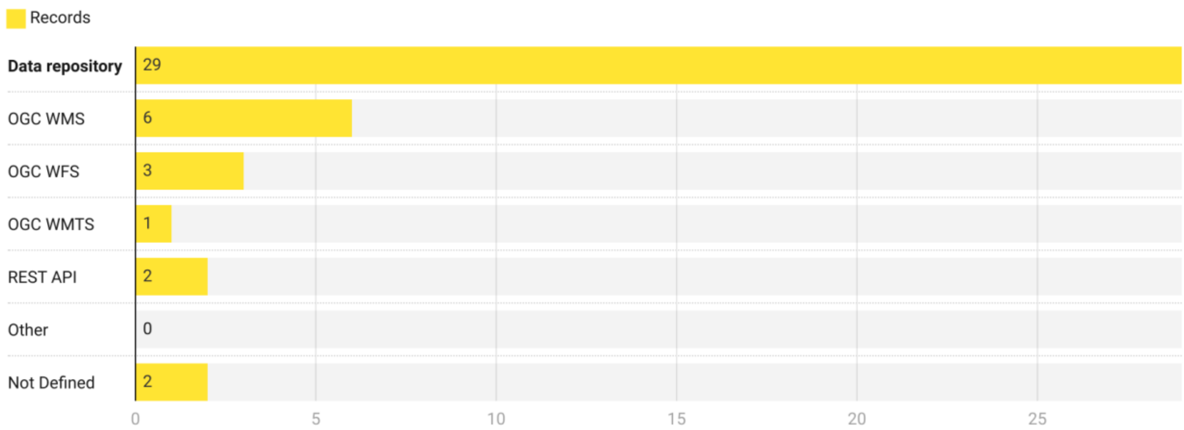



Figure 14: Chosen interfaces (User or Machine-readable), under which the data is available.

Concluding this analysis with the two remaining objectives under examination, we shall comment that the Riga's data ecosystem depicts a different picture, as the data observations have made their early appearance from the 1970s'. Nevertheless, a similar digital growth seems also to characterize this city in the last decade. Additionally, several annual products are generated as aggregated results of these parameters or for the EO-driven products, with fewer data assets to be updated with a coarser frequency. Finally, in this city as well, we have 11 datasets that either are not updated, or no information was given.

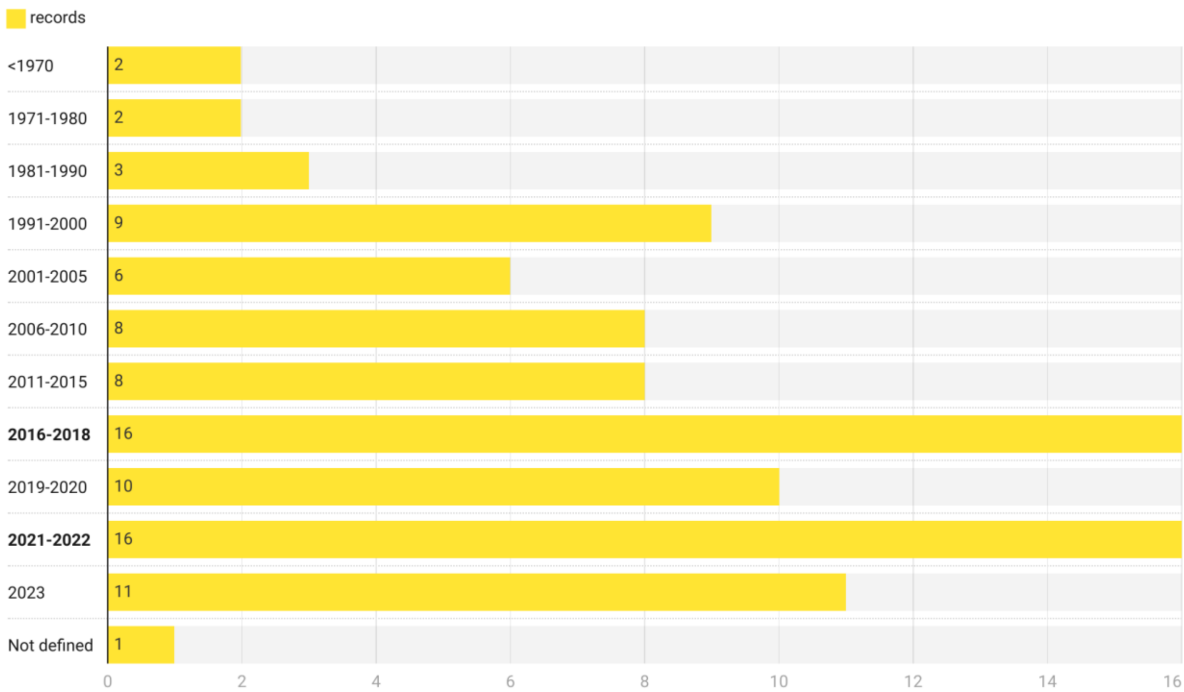



Figure 15: Distribution of the number of data sources per the user-determined time intervals.

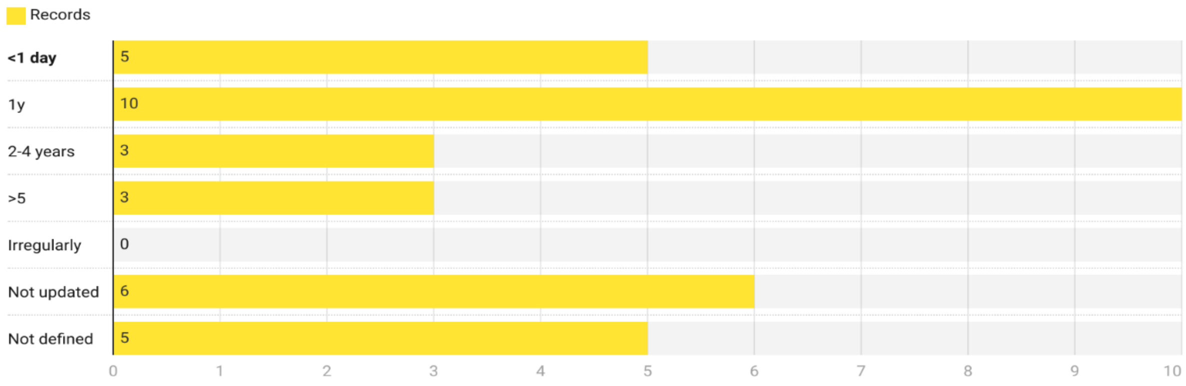
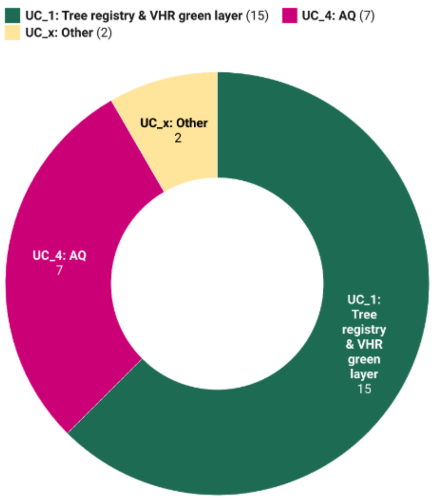



Figure 16: Distribution of the number of data sources per the user-determined classes of update

## 5.2.3 Data Ecosystem Mapping of Athens

The City of Athens has nearly 700.000 residents[48], and is part of a 3.75 million people metropolis, which has during the past years been experiencing a multivariate of crises that affect financial stability and the environment. Since 2016, the city of Athens has joined the global initiative of 100 Resilient Cities (RC100), having the ambition to form meaningful urban resilience through plans that account for the city's entire urban ecosystem. It addresses not only the city's shocks, such as earthquakes, and civil unrest, but its chronic stresses, such as heat waves, homelessness, poor air quality, long-term unemployment, and other continuous pressures that erode the urban fabric. The 2030 Resilience strategy of Athens was formulated to identify and prioritize the aforementioned aspects, and several monitoring initiatives have been established. This perspective can be supported from our findings, illustrated in Figure 17.




| UCs | Data assets |
|---|---|
| UC_1 | Tree registry [1st, 2nd, 3rd, 4th, 5th, 6th, and 7th] districts of Athens; trees (single records) National garden of Athens and maps; Tree records reported as pavement objects (Athens district and historical center); Records with plant boxes and parterre; |
| UC_4 | National Air Pollution Monitoring Network (e.g. PM10, CO, NO, NO2, O3, SO2); PANACEA PM2.5 Sensor Infrastructure (Athens Network); |
| UC_x | Noise volume (Lden index); Pavements |

Figure 17: (Left-side): Number of datasets per Use Case; (Right-side): Brief description of the identified datasets classified in each category.

Such initiatives are predominately organized by the Municipality of Athens, (Figure 18 and Table 12) and the 7 operating greening departments, which are keen on mapping the locations and species of the existing trees. This initial effort was the triggering point for the National Garden to provide the first geospatial layer that presents the trees records that have been

[48] https://resilientcitiesnetwork.org/downloadable_resources/Network/Athens-Resilience-Strategy-English.pdf

cultivated or in general exist within the premises of the National Garden. An unexpected addition to the aforementioned was that through this data review process, we managed to identify six additional geospatial datasets, which are normally characterized as "pavements' obstacles". The majority of these obstacles are trees or low-height plants stored in parterres or plant boxes.

Furthermore, two organizations are responsible to provide valuable observations of air pollutants, e.g. PM2.5 and 10, CO, NO, NO2, O3, and SO2 with the first to be pertained under the jurisdiction of the Municipality of Athens and the second to be as an outcome the PANhellenic infrastructure for Atmospheric Composition and climate change (PANACEA)[49] consortium. Finally, noise measurements are generated by the Hellenic Ministry of Environment and Energy (YPEN).

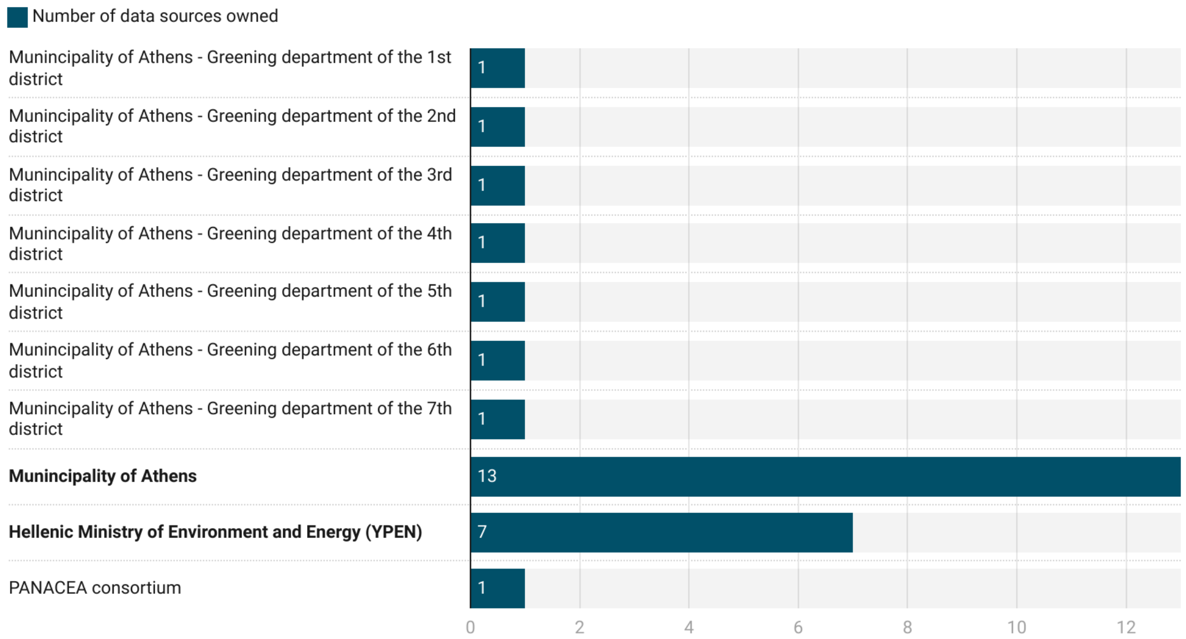



Figure 18: Number of datasets per identified data actors of Athens.

Table 12: Datasets of the different identified data actors of Athens.

| Data owners | Data assets |
|---|---|
| Greening department of the 1st district | Tree registry 1st district of Athens |
| Greening department of the 2nd district | Tree registry 2nd district of Athens |
| Greening department of the 3rd district | Tree registry 3rd district of Athens |
| Greening department of the 4th district | Tree registry 4th district of Athens |
| Greening department of the 5th district | Tree registry 5th district of Athens |
| Greening department of the 6th district | Tree registry 6th district of Athens |
| Greening department of the 7th district | Tree registry 7th district of Athens |
| Municipality of Athens | Tree registry [1st, 2nd, 3rd, 4th, 5th, 6th, and 7th] districts of Athens; trees (single records) National garden of Athens and maps; Tree records reported as pavement objects |

[49] https://panacea-ri.gr/?lang=en

| | |
|---|---|
| | (Athens district and historical center); Records with plant boxes and parterre; National Air Pollution Monitoring Network (e.g. PM10, CO, NO, NO2, O3, SO2); Noise volume (Lden index); Pavements |
| Hellenic Ministry of Environment and Energy (YPEN) | National Air Pollution Monitoring Network (e.g. PM10, CO, NO, NO2, O3, SO2); |
| PANACEA consortium | PANACEA PM2.5 Sensor Infrastructure (Athens Network); |

Analysing the data formats (Figure 19) and interfaces (Figure 20), the majority of the collected observations are stored in statistical spreadsheets, or CSV data formats with the alternative option of the ASCII to be given only once. The tree geospatial layers are predominantly point-wise vectorized records and are by default stored in both ESRI shapefile and GeoJSON data formats. Five data assets are offered in non-editable formats, e.g. the Adobe Reader Portable Document Format (PDF) and the Joint Photographic Experts Group (JPEG)/Portable Network Graphics (PNG). Finally, the majority of the aforementioned records are categorized as “closed datasets” according to the Data spectrum (see chapter 6) and thus are not stored in internal data repositories, or have APIs for sharing. Only two records have been declared to be provided as OGC-compliant services.

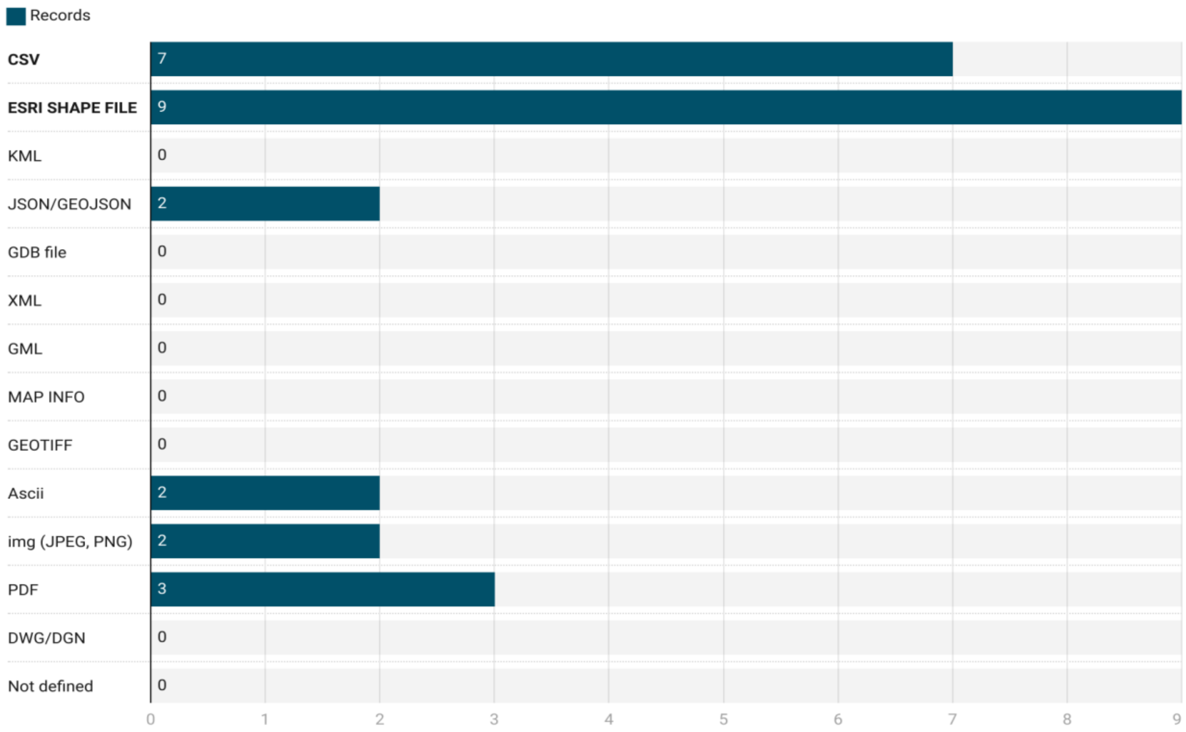



Figure 19: Chosen data format to store the collected data assets.

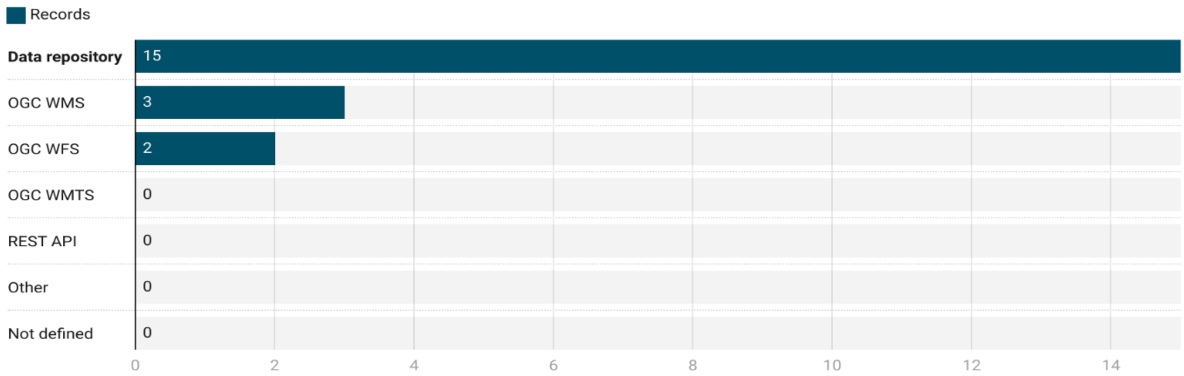



Figure 20: Chosen interfaces (User or Machine-readable), under which the data is available.

Finally, it appears that the last decade was pivotal in the digitalisation process of the city of Athens since all the measured datasets are mostly collected since 2015, (Figure 21), and only the AQ data with some consistency. For the rest 10 data sources, additional investigation is required, (Figure 22).

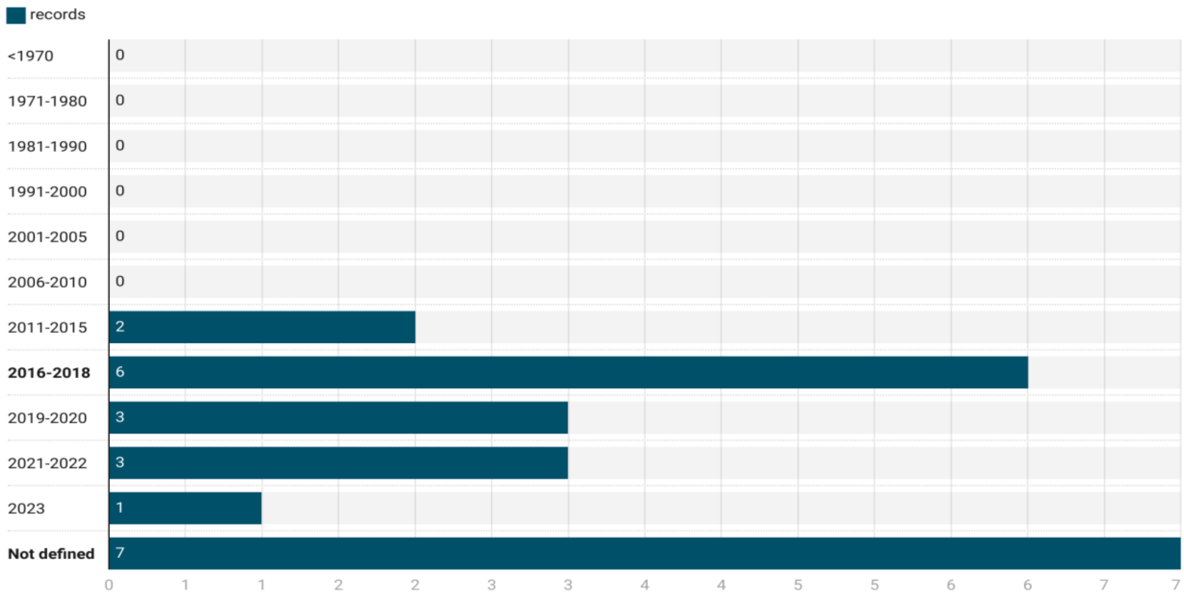



Figure 21: Distribution of the number of data sources per the user-determined time intervals.

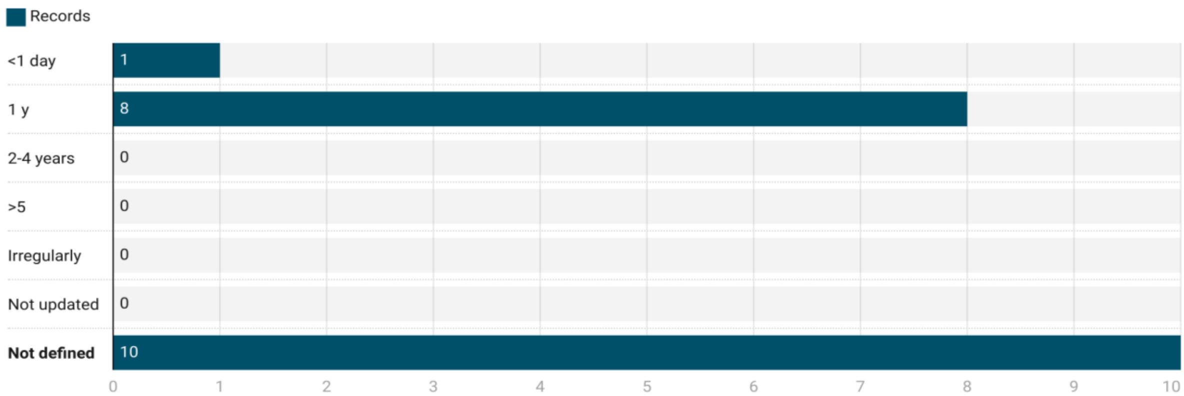


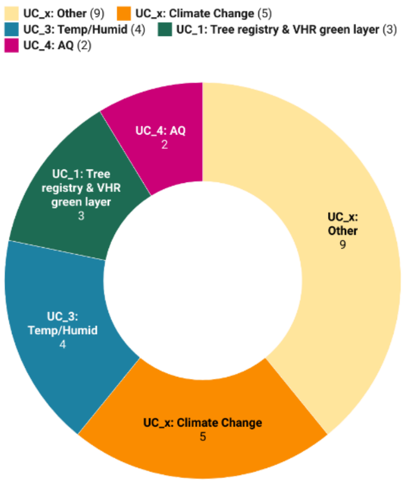

Figure 22: Distribution of the number of data sources per the user-determined classes of update frequency

### 5.2.4 Data Ecosystem Mapping of Mannheim

Mannheim is a strong advocate that the process of digitalization is of paramount importance to address climate change and protect the natural resources, and other challenges that humanity will face[50]. These formulated the basis of the Mission Statement Mannheim 2030SDG strategy[51]; that contains tangible outcomes and measures, which envision the future of Mannheim. The mission statement of this strategy is summarised in the following six objectives:

- Attempt to describe what life will be like in the coming years in Mannheim and what it means to be a Mannheim resident, and after all which are the challenges and prerequisites to achieve an adequate QoL.
- A digital tool to enable the city community to get involved in the future of Mannheim in a meaningful way and understand what the residents will experience.
- Indicators for approaching a series of measures and comprehending with direct visualisations the impact of the regulated achievements in the short term, which will update and reshape the deployment of long-term goals.
- Avoid creating a plan that replaces all other strategies, and be keen on harmonizing the subsequent strategies upon a common wider mission
- Particularly effective when the city community commits to consistent action in accordance with the 17 sustainability goals.
- Revitalize bureaucratic procedures and promote innovation in the city.

Upon these tangible goals, the city of Mannheim has been devoted to this digitalization process and constant monitoring of several aspects of essential interest to measure the sustainability of the city. This outcome can be also visualised in Figure 23.




| UCs | Data assets |
|---|---|
| UC_1 | Aerial 3D laser scanning survey; Public green spaces; Newly planted trees |
| UC_3 | Surface infrared temperature; Isothermal map; temperature, humidity (2 sources), |
| UC_4 | AQ [NOx, CO2] |
| UC_x | Wind directions, cold air corridors, Heat areas; Buildings and land sharp GIS data; Noise data; Soil moisture and soil temperature; Heavy rain hazard maps; Precipitation; Wind speed and Direction; Cool places maps; Air pressure; AQ2 [precipitation, global radiation, wind] |

Figure 23: (Left-side): Number of datasets per Use Case; (Right-side): Brief description of the identified datasets classified in each category

The city of Mannheim, which could be characterised as the key player in the data-gathering process (Figure 24 and Table 13) has generated the open geoportal of Mannheim, namely "StadtMannheim", in which several geospatial features exist, focused on tree monitoring and preservation points, as well as air quality and climate change. This effort is intensively supported by two additional organisations of the city, the Smart City of Mannheim and the

[50] https://www.mannheim.de/sites/default/files/2019-03/Mission%20Statement%20Mannheim%202030_%2013.03.2019_English_WebFile.pdf

[51] https://www.local2030.org/pdf/vlr/mannheim-vlr-2020.pdf

State Institute for the Environment of Baden Württemberg. This specific platform stores data assets related to the locations of trees dated up to 2014 and the distribution of public green spaces. Complementary to these, a rather dense sensor network of about 45 sensors seems to exist within the city, collecting meteorological and air quality data, and this network is to be sincerely expanded by the end of 2023, deploying 460 additional sensors. Further details are given below.

- **12 sensors** collecting **precipitation** observations and **20 additional sensors** to be planned to be included.
- **30 sensors** focused on the **wind speed and direction measurements**, with **400** to be about to come by the end of 2023.
- **3 stations** measure the variables of precipitation, temperature, global radiation, humidity, and wind.
- **2 air quality sensors**, collecting observations related to NOx and $CO_2$ pollutants, as well as other complementary observations, such as temperature, humidity, and air pressure.
- **40 surface temperature sensors** are planned to be installed, measuring the surface thermal status on a tram and thus investigating the impact of green spaces. Additionally, temperature, humidity and pressure parameters will be as well collected.

Geospatial rasterised products have been produced and disseminated openly, showcasing the isothermal profile over the city, heat-stressed areas, and on the contrary the existing cooling places. Furthermore, two supplementary soil sensors measuring the soil temperature and moisture are under schedule to be installed. Finally, Mannheim explores other important elements that are related to noise quality and the susceptibility of the city to floods. 3D laser scanning surveys have been scheduled to be conducted by the end of the year and could be disseminated within the consortium of UR upon request, which could assist the identification of trees that are located in private lands.

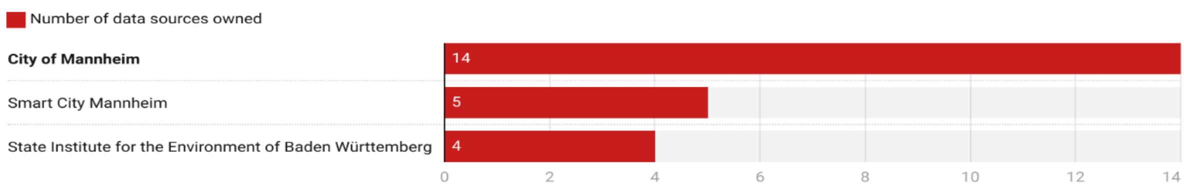



Figure 24: Number of datasets per identified data actors of Mannheim

Table 13: Datasets of the different identified data actors of Mannheim.

| Data owners | Data assets |
|---|---|
| City of Mannheim | Aerial survey; Public green spaces; Newly planted trees; Isothermal maps; Wind directions, cold air corridors, Heat areas; Buildings and land sharp GIS data; Noise data; Heavy rain hazard map; Cool places map |
| Smart City Mannheim | Surface infrared temperature; Soil moisture and soil temperature; Precipitation; Wind |
| State Institute for the Environment of Baden Württemberg | Air measurements 2 stations: NOx, CO2, temperature, Humidity, air pressure |

Majority of the detected data sources are available as ESRI shapefiles, even in the case of the geospatial maps that cover the whole city (e.g. noise maps, heavy rain hazard maps, cool places, etc.). A single data asset related to the 3D laser scanning surveys seems to be created in GeoTIFF form, whereas since this specific data asset is closed a further investigation is necessary to validate our assumption. Rather than that, several data sources, which contain meteorological and air quality measurements are stored in CSV formats. Finally, a single data source is classified as “Not defined”, since the deployment of the soil moisture and temperature sensors are still in the experimental phase.

Majority (14 records) of data assets are accessible through the open data portal of the city of Mannheim. Only for two options the RESTful API is provided as an alternative of data access.

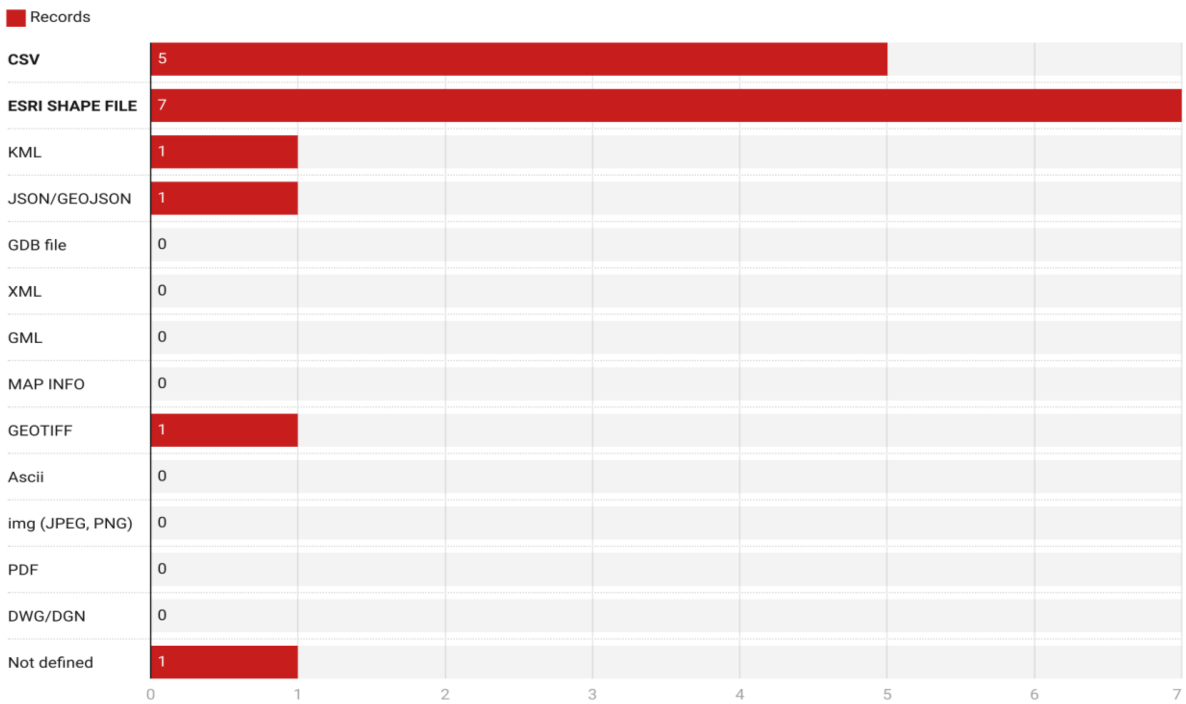



Figure 25: Chosen data format to store the collected data assets.

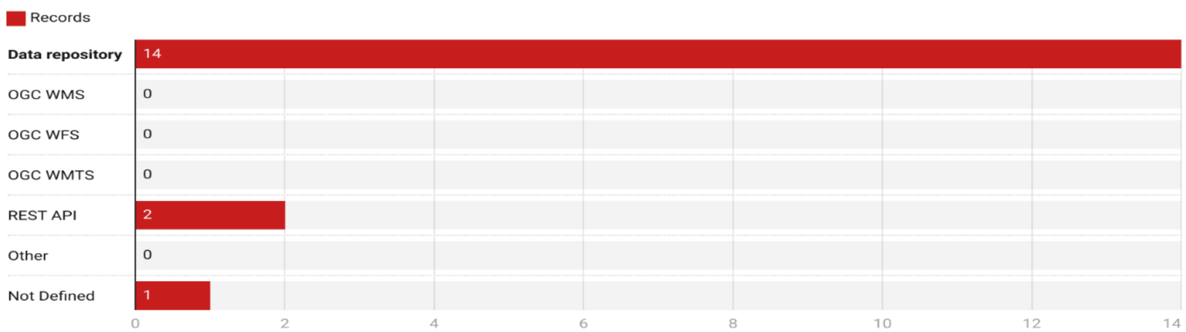



Figure 26: Chosen interfaces (User or Machine-readable), under which the data is available.

Mannheim reveals a similar picture with Athens, as the revolution of digitalisation has started within the last 20 years. In particular, the time intervals of 2011-2015 and 2021-2022 are described as those of great interest in producing information valuable to the city.

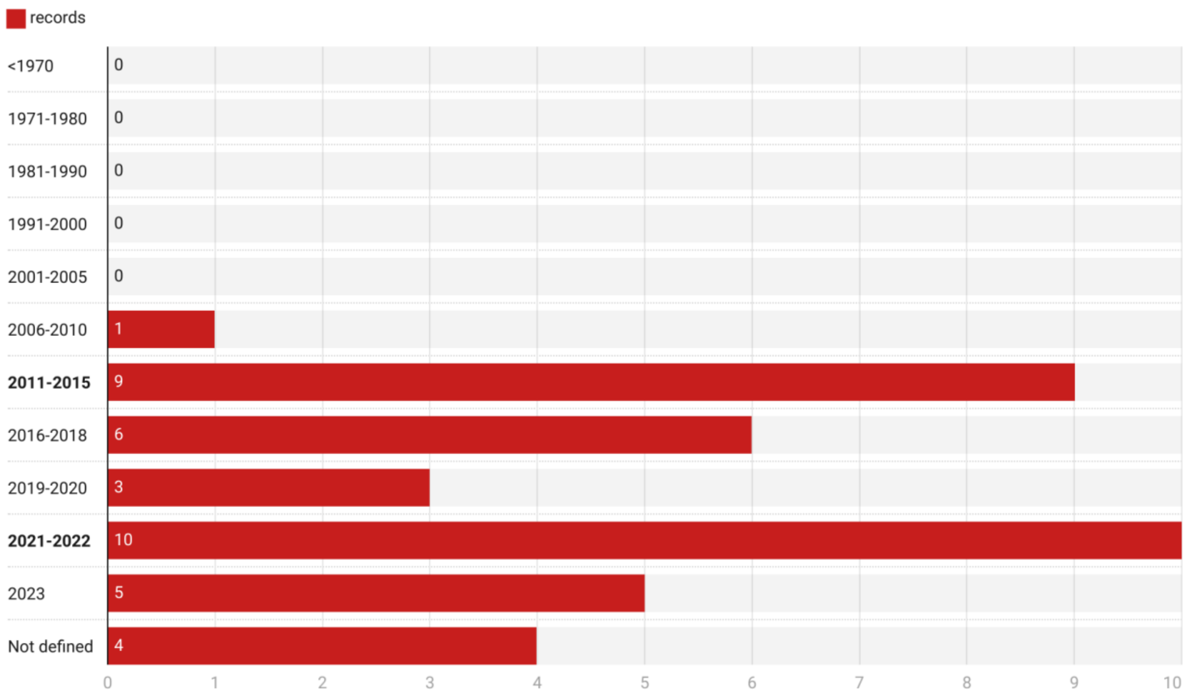



Figure 27: Chosen interfaces (User or Machine-readable), under which the data is available.

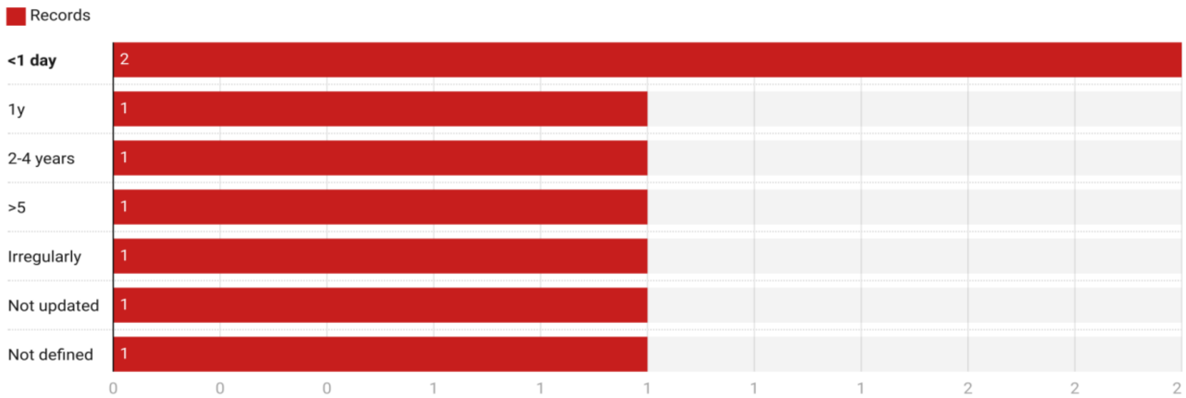


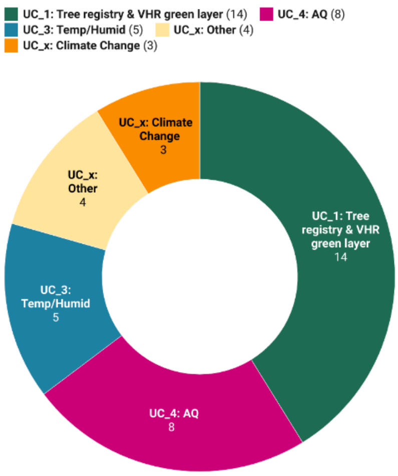

Figure 28: Distribution of the number of data sources per the user-determined classes of update frequency

## 5.2.5 Data Ecosystem Mapping of Cascais

Cascais is severely affected by intense flash floods, which are expected escalate causing devastating results in the urban regions of the city, as well as buildings of significant importance (e.g., historical buildings). Considering the above, Cascais was the first Portuguese municipality to devote itself to the UN-SDG goals. In 2009 it developed its first Strategic Plan for Climate Change Adaptation Action (PECAC)[52], a risk management, adaptation and mitigation framework that is keen to be continuously reformulated based on the current trends and necessities. In this vision, it is advocated that the most recent measures include the creation of green spaces and corridors in urban areas, beach and shore face

[52] https://www.lneg.pt/wp-content/uploads/2022/05/Cascais-prioritEE-Plus-29-abril.pdf

nourishment, rehabilitation and restoration of rivers, establishment and restoration of riparian buffers and awareness citizen science campaigns for behavioural change and engagement[53].




| UCs | Data assets |
|---|---|
| UC_1 | Green urban areas; Orthophoto/VHR layers; Protected trees; Land Use and Occupation Charter; Tree registry; Interventions in Trees; Potential Natural Vegetation; Green urban areas; Assessment of ecosystem services [carbon removal, pluvial runoff avoided, pollutant removal]; Green urban areas |
| UC_3 | 5 stations [Humidity, Temperature]; 11 AQ stations [Humidity; Temperature] |
| UC_4 | 11 AQ stations [NO2, O3, PM2.5, PM10, SO2, CO]; Predicting air Quality for Outdoor Activities |
| UC_x | 11 AQ stations [noise, pressure, wind speed and direction, precipitation]; 5 stations [Barometer, Evapotranspiration, Precipitation, Wind Speed, Direction, Max, WindSamps, Solar Radiation, UV]; Climate Change Adaptation Plan; Fire hazard map; Fire risk map |

Figure 29: (Left-side): Number of datasets per Use Case; (Right-side): Brief description of the identified datasets classified in each category.

Measurable results are required to address these mitigation policies. More specifically, as shown in Figure 29, several parameters are covered related to the existence of urban greening spaces, as well as preservation practices, monitoring their conditions and interventions that may have taken place. The Cascais Town Hall local authority in collaboration with the dedicated department of the Green Infrastructure Management Division has been responsible for creating and monthly updating a tree registry in urban space and therefore providing information on the location, characteristics, vegetation species and their scientific name, whereas by far more crucial statements of interventions that might have taken place. In collaboration with the i-Tree project, essential variables to the ecosystem status and the benefits of the presence of the greening resources are provided, with the following three parameters 1) carbon removal; 2) pluvial runoff avoided; and 3) pollutant removal to be declared as indicative examples of their research. Furthermore, expanding their practices on a local scale, subsequent observations of the existing green urban areas, such as the coverage of municipality parks and gardens (i.e., location, characteristics, responsible owner, and resources integrated), are provided as well.

The Directorate-General for Territory and the Municipal Environment Company of Cascais (EMAC)have a keen interest to contribute to the detailed monitoring of the green spaces of Cascais, They add to the data ecosystem with VHR orthophoto images over Cascais (1995-2022), land cover geospatial datasets, and data from the “Terras de Cascais” project. Lastly, the institute for Nature Conservation and Forests of Cascais provides a geospatial layer with locations of natural greening resources that are in need of special protection.

Continuing, Cascais cis a city that has a major interest in aspects related to AQ and climate resilience. A joined collaboration between EMAC and a private organisation, namely QART (Monitoring and Urban Environmental Mapping Solutions) is taking place, to operate 16 in-situ sensors, which are responsible to monitor several meteorological and AQ variables (Figure 29). Subsequently, the collected PM2.5 observations had been further processed in terms of



[53] https://base-adaptation.eu/participatory-review-strategic-adaptation-plan-cascais-portugal.html

the World Data League (WDL) hackathon, utilising AI models, to generate predictive measurements on air quality. Finally, the Forest Technical Unit of the Cascais Town Hall is involved in the generation of fire hazard and risk maps.

Distribution of data assets based on owners are elaborated in Figure 30 and Table 11.

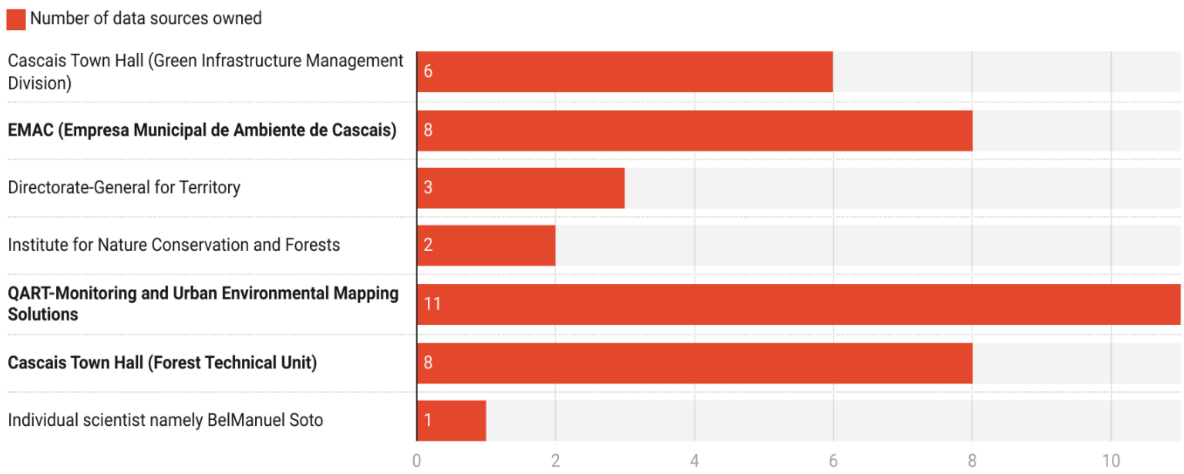



Figure 30: Number of datasets per identified data actor of Cascais

Table 14: Datasets of the different identified data actors of Cascais.

| Data owners | Data assets |
|---|---|
| Cascais Town Hall (Green Infrastructure Management Division) | Tree registry; Interventions in Trees; Assessment of ecosystem services of urban trees; Green urban areas |
| EMAC (Empresa Municipal de Ambiente de Cascais) | Green urban areas; Potential Natural Vegetation; Green urban areas; 5 meteorological stations [Barometer Evapotranspiration, Humidity, Precipitation, Temperature Wind Speed, Direction, Max and WindSamps, Solar Radiation, UV] |
| Directorate-General for Territory | Orthophoto or VHR layers of Cascais; Land Use and Occupation Charter |
| Institute for Nature Conservation and Forests | Protected trees |
| QART-Monitoring and Urban Environmental Mapping Solutions | 11 AQ stations [NO2, O3, PM2.5, PM10, SO2, CO, noise, pressure, relative Humidity, temperature, Wind speed and direction, Precipitation, Classification of air quality] |
| Cascais Town Hall (Forest Technical Unit) | Fire hazard map; Fire risk map |
| Individual scientists namely BelManuel Soto | Predicting air Quality for Outdoor Activities |

Moving towards the additional aspects related to the commonly adopted data formats (Figure 31) and interfaces (Figure 32) that are of use in Cascais, we could see that by far CSV is of the most favourable data formats, commonly adopted for meteorological and climatological observations. In addition, it seems that 7 records are declared to hold the geospatial locations of the monitored observations, rather than the thematic information. For these data assets, the rather known ESRI shapefile is used and therefore complemented with the KML and JSON or GeoJSON data formats. A single data asset, which is assumed to be the orthophoto images, is provided in GeoTIFF format. Finally, the AQ data sources generated by QART are disseminated through the public upon a developed closed-based API, in which the observations along with the corresponding metadata are given in XML format.

Concluding with the technological aspects of this analysis, in Figure 32, it is observed that the majority of the aforementioned data sources are disseminated via the open data platform of Cascais, with 3 records to be able to be subsequently explored in a more "machine-readable" manner, using the OGC-compliant services of WMS, WFS and WMTS, with the latest to be adopted for the image datasets. For a single data asset, we weren't able to find information either for the data format or the interface under which is provided.

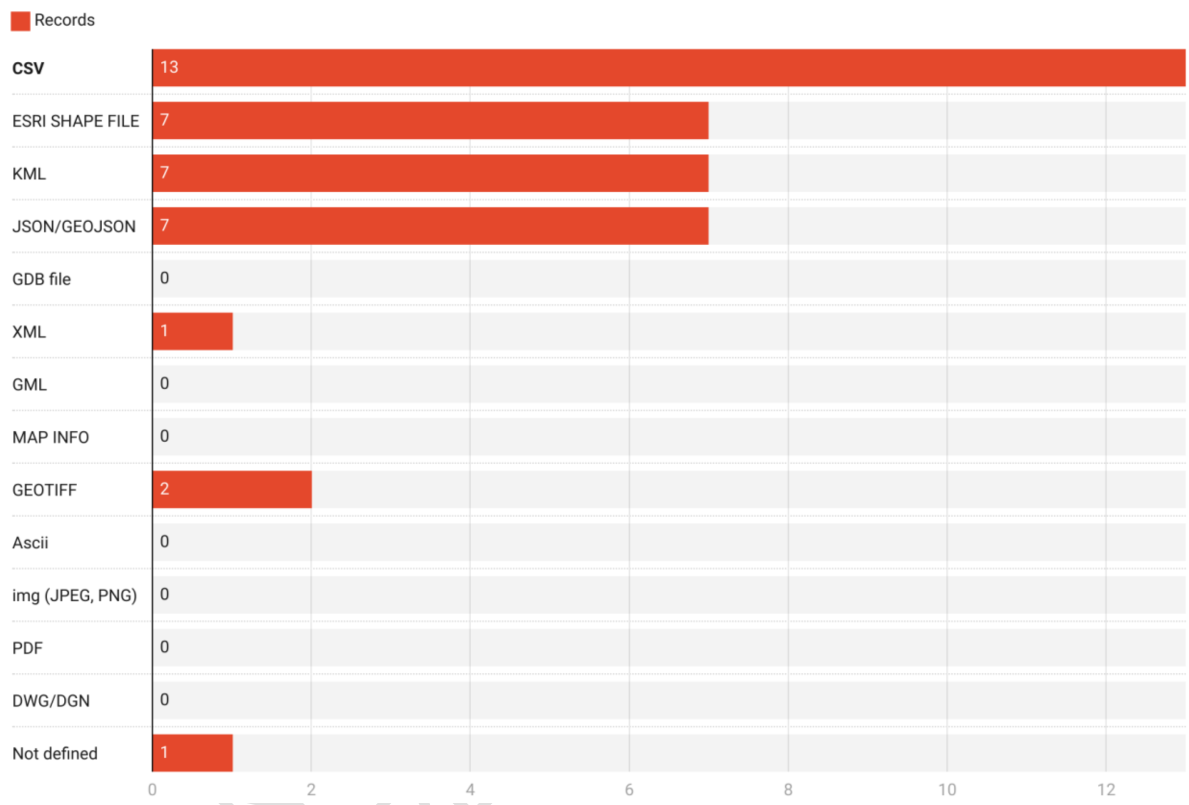



Figure 31: Chosen data format to store the collected data assets

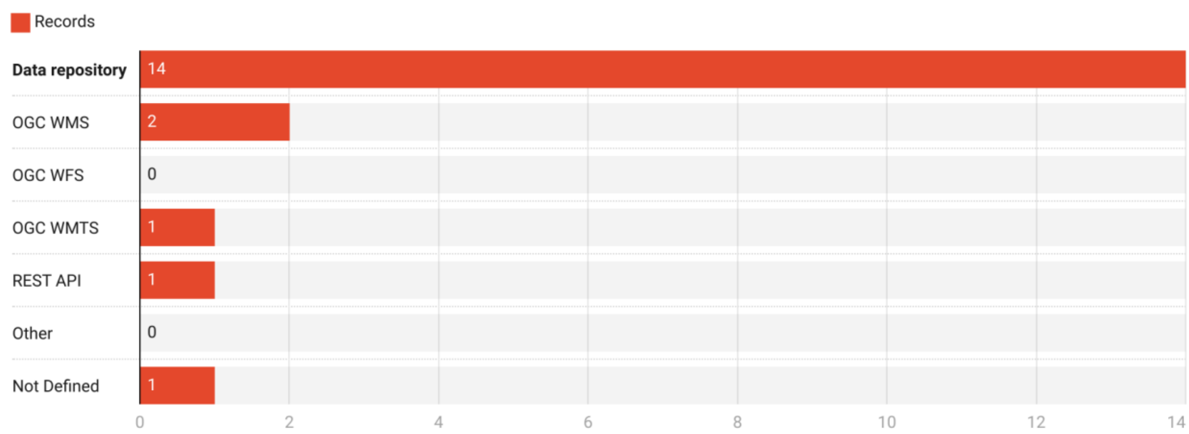



Figure 32: Chosen interfaces (User or Machine-readable), under which the data is available.

Examining the tendency of the city to perpetually monitor its sustainability profile, it appears that this initiative has started since the 80s', when the first aerial surveys took place to provide VHR-RGB maps; an effort that maintains until now. This digital evolution seems to subsequently expand since 2000 and onwards. Nonetheless, during the last two years, the DEM of Cascais can be characterised by a spiralling increase, holding the most of the identified data observations. However, the majority of the identified data assets are not characterised by a specific data update planning. In particular, the data updates are commonly described by irregular updates. An exception is the climatological and meteorological datasets, which are provided both as daily observations and annual aggregated and calibrated datasets.

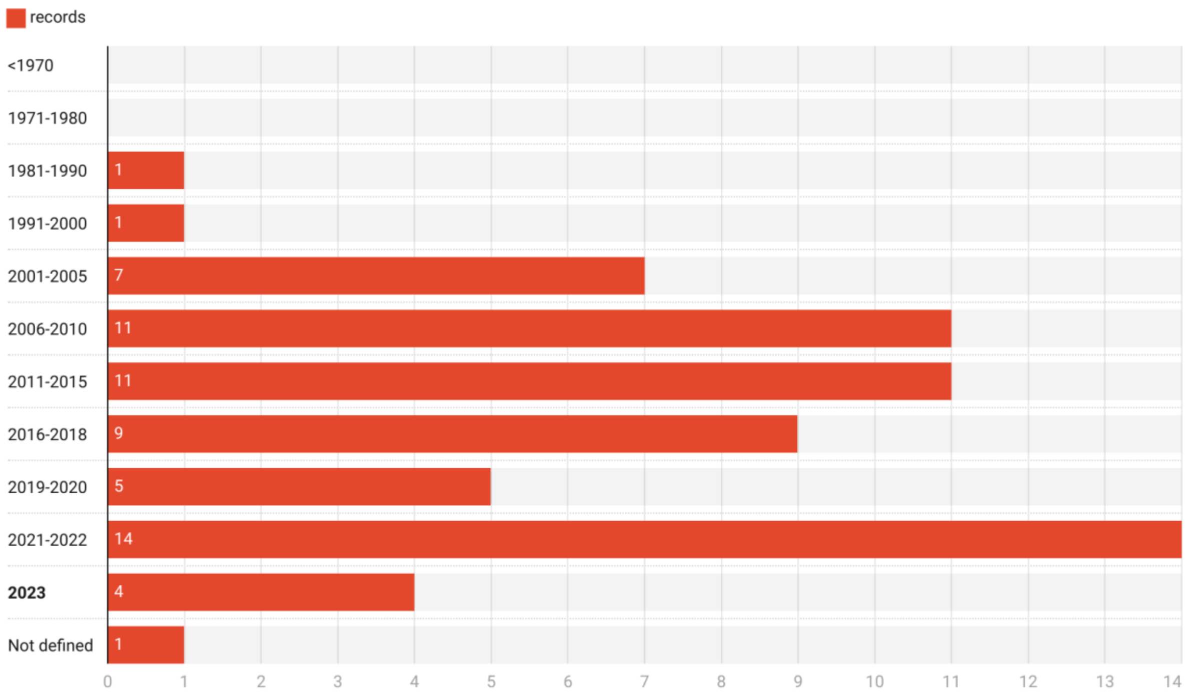



Figure 33: Distribution of the number of data sources per the user-determined time intervals.

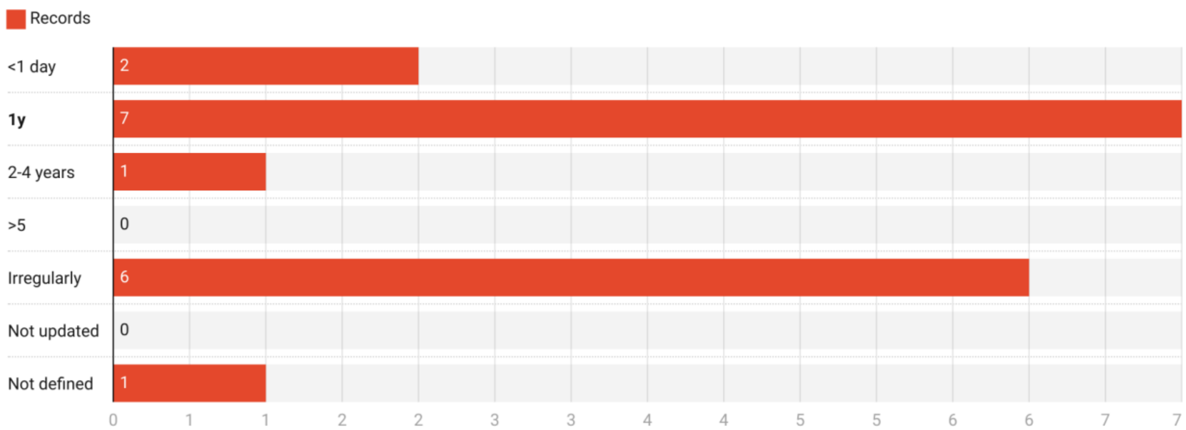
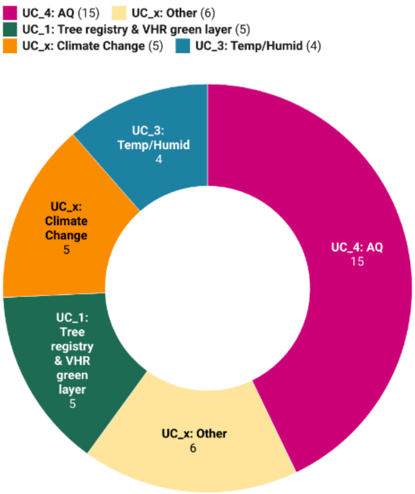



Figure 34: Distribution of the number of data sources per the user-determined classes of update frequency

### 5.2.6 Data Ecosystem Mapping of Utrecht

Utrecht formulated the 2040 Spatial Strategy, confirming t its commitment to the fulfilment of the 17 SDGs, and hence achieving a healthy balance between densification and greening. Some indicative objectives of this strategy are an expansion of the greenery in the city by 440 hectares with 60,000 trees, and the substantial improvement of the air conditions. To achieve these commitments, the municipality of Utrecht founded the local foundation of Utrecht4GlobalGoals[54], highlighting the importance of coalition as well between the key drivers of the city[55].

To support this vision, a plurality of stakeholders in Utrecht have made substantial efforts to monitor key aspects of the city, in order to better facilitate the efforts of this sustainable rejuvenation. As a result, several geospatial datasets have been generated, revealing the greening and air quality conditions of the city, and therefore its sustainability level. They are briefly illustrated in Figure 35, and elaborated below.




| UCs | Data assets |
|---|---|
| UC_1 | Tree map (species, year, tree number); Old trees 80 years and older; Forest plants with trees; Green structure vision map; Land use |
| UC_3 | Snifferbike (Temperature) aggregated values in hexagons; Green Bicycle roots (temperature, humidity); Measure Together-CS [Temperature] |
| UC_4 | Snifferbike (PM2.5); and aggregated in hexagons; Green Bicycle roots (PM2.5, PM10, PM1); Annual Mean PM2.5; Annual Mean PM10; AQ1 [NO2, PM2.5, PM10]; AQ2-68 observation points [CO2, NO, NO2, PM10, O3]; Measure Together-CS [PM2.5, PM10, NO2, NH3] |
| UC_x | Green Bicycle roots (air pressure); Traffic light intensity; Traffic data (bike, car, truck, pedestrian); Noise contours (Lden); Measure Together-CS [Noise and Water Quality]; Climate change maps (Number of tropical days, of summery days, warm days); Warm nights; Heat maps (2050 Best and Worse-case CC scenario); Risk drought stress; Drought susceptibility nature; UHI |

Figure 35: (Left-side): Number of datasets per Use Case; (Right-side): Brief description of the identified datasets classified in each category.

Having a quick look at the distribution of the data assets alongside the data owners of the city (Figure 36, Table 11), a significant interest is foreseen from the municipality of Utrecht in the collection of information that is related to the monitoring of urban green areas. More specifically, several detailed datasets including a tree registry exists. The data collections also cover the tree records in forested areas. Moreover, dedicated efforts are given to map aged trees, indicating their ambition of initiating a particular planning of preservation. All these sources comprise the green structure vision map of Utrecht. The map is an outcome of the "2018 Green Structure Plan Update" initiative, which intends to identify the greening areas in the city, preserve and substantially improve them. Lastly, the land cover dataset aims to thoroughly present the wider coverage of the land categories that exist over the city.

Several stakeholders e.g. public authorities: the Province of Utrecht and the department of Traffic Systems, the National Institute of Health and Environment (RIVM), the Central

[54] https://utrecht4globalgoals.nl/
[55] https://use.metropolis.org/case-studies/utrecht--a-global-goals-city#casestudydetail

Instrument of Monitoring the Air Quality (CIMLK), and the Royal Netherlands Meteorological Institute (KNMI), and private firms: Civity and SODAQ, are contributing to the deployment of conventional and low-cost sensors from the Snifferbike, Measure together and Green Bicycle roots projects, and provide measurements on be PM2.5, PM10, PM1, CO2, NO, NO2, PM10, O3, etc. Also,  meteorological observations are also collected. Subsequently, three additional parameters seem to be of essential interest to the city, namely noise observations, the quality of the water's natural resources and the pollution caused by traffic lights.

Eventually, the general ambition of Utrecht to be at the forefront and "fight" against climate change, has been a triggering point of generating the Climate Impact Atlas project, which disseminates through a web mapping interface, the data sources themselves and intuitive narrative stories., The results of the best-case and worst-case climate change 2050 scenarios in several domains that impact our lives and the surrounding area are illustrated. The models' results of flooding, waterlogging, drought, and heat (UHI and heat maps), as well as the susceptibility of the area in drought, are showcased following the same climate impact categorisation as the Delta Plan on Spatial Adaptation. All the aforementioned findings are a result of the following four research institutions, i.e. KNMI, Wageningen Environmental Research, KWR Water Research Institute, Atlas Natural Capital, and the Witteveen+Bos**,** which belongs to the private sector.

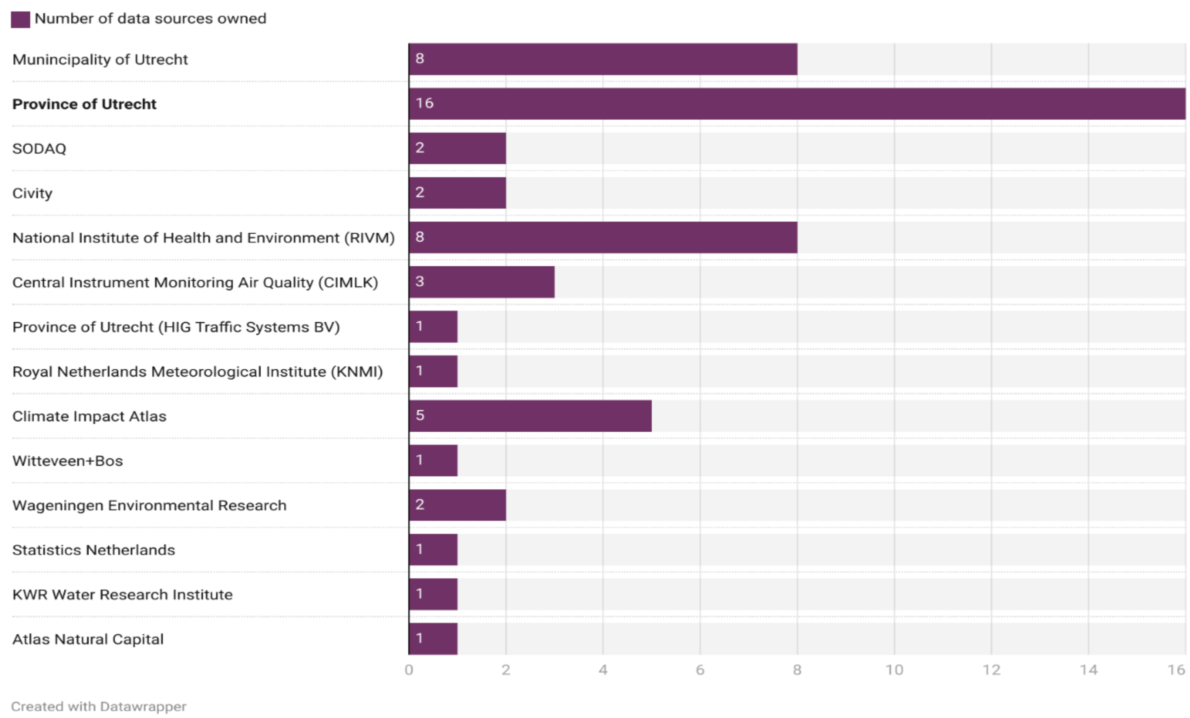



Figure 36: Number of datasets per identified data actor of Utrecht.

Table 15: Datasets of the different identified data actors of Utrecht.

| Data owners | Data assets |
|---|---|
| Municipality of Utrecht | Tree map (species, year, tree number); Old trees 80 years and older; Forest plants with trees; Green structure vision map |
| Province of Utrecht | Green Bicycle roots [temperature, humidity, air pressure, PM2.5, PM10, PM1]; Annual Mean PM2.5 and PM10 concentration; |

| | |
|---|---|
| | Snifferbike [PM2.5 and Temperature]; GES noise contours (Lden); Traffic data (bike, car, truck, pedestrian) |
| SODAQ | Snifferbike [PM2.5 and Temperature] |
| Civity | Snifferbike [PM2.5 and Temperature] |
| National Institute of Health and Environment (RIVM) | Measure Together-CS Noise, PM2.5, PM10, NO2, NH3, Temperature and Water Quality; Snifferbike [PM2.5 and Temperature] |
| Central Instrument Monitoring Air Quality (CIMLK) | AQ1 [NO2, PM2.5, PM10]; Traffic light intensity; AQ2-68 observation points [CO2, NO, NO2, PM10, O3] |
| Province of Utrecht (HIG Traffic Systems BV) | Traffic data (bike, car, truck, pedestrian) |
| Royal Netherlands Meteorological Institute (KNMI) | Climate change maps (Number of tropical days, summery days, warm days); Warm nights; Heat maps (2050 Best and Worse-case climate change scenario); Risk drought stress; Drought susceptibility nature; UHI |
| Climate Impact Atlas | Land use |
| Witteveen+Bos | Heat maps (2050 Best and Worse-case climate change scenario) |
| Wageningen Environmental Research | Warm nights; Heat maps (2050 Best and Worse-case climage change scenario) |
| Statistics Netherlands | Land use |
| KWR Water Research Institute | Risk drought stress |
| Atlas Natural Capital | UHI |

The commonly adopted data formats (Figure 37) and interfaces (Figure 38) of Utrecht are highlighted here.  We see two main types of data formats: the first is the very known CSV format predominately presented in meteorological and climatological data sources and the ESRI shapefile for the geospatial observations. In limited observations, the KML and the GeoJSON data formats are complementary and provided as an additional source to accompany the dominant geospatial format. Moreover, as was expected, in this city as well, the majority of the detected data assets are publicly available through the different repositories that each data owner has deployed. Lastly, it appears that on 25 occasions, alternative access end-points can also be utilised by users, which could be described as more “machine-friendly”. For the larger-scale datasets OGC-compliant WMS services and APIs are available, whereas for the local-scale vectorised observations the WFS has been chosen.

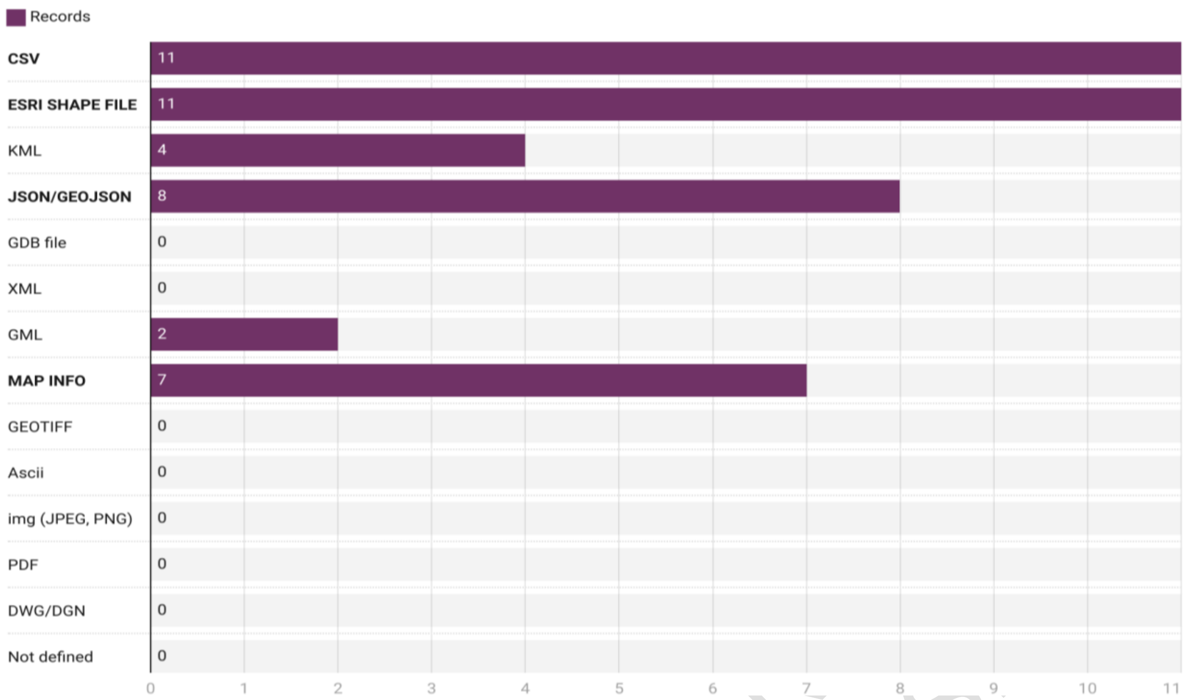



Figure 37: Chosen data format to store the collected data assets.

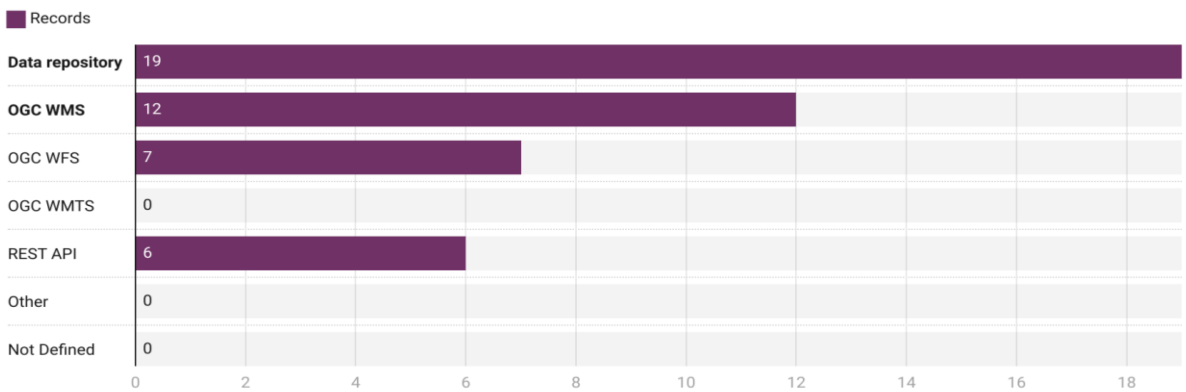



Figure 38: Chosen interfaces (User or Machine-readable), under which the data is available.

Our findings illustrated in Figure 39, show a significant environmental monitoring efforts in Utrecht in the last 10 years. In the case of Utrecht as in the previous pilots, the majority of the identified data assets seems to be generated in 2021-2022 (Figure 40).

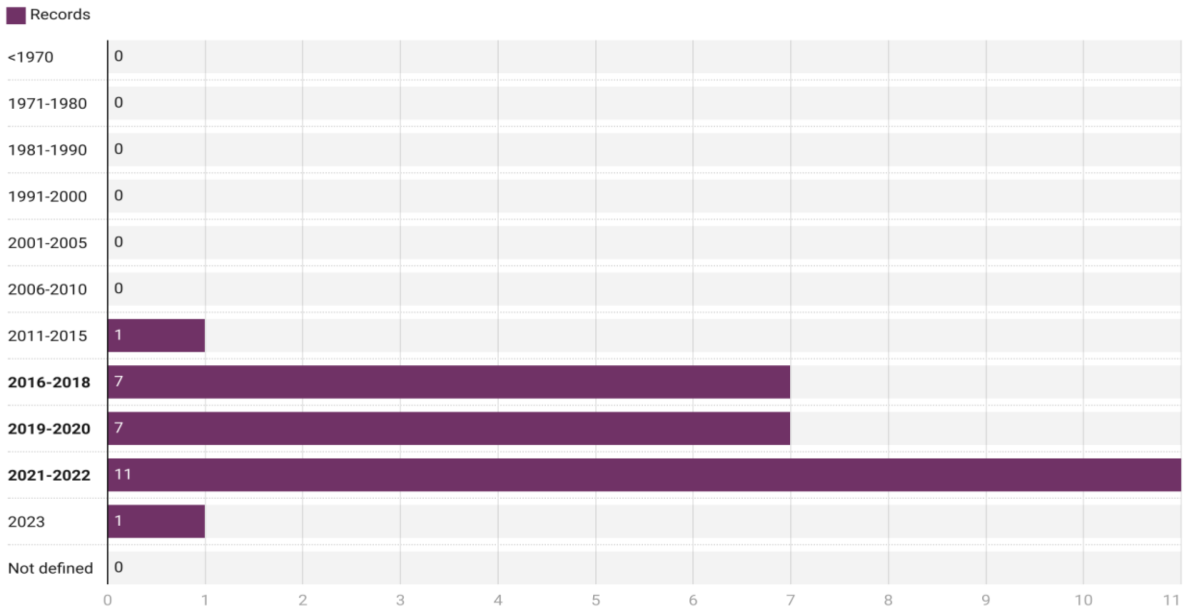



Figure 39: Distribution of the number of data sources per the user-determined time intervals.

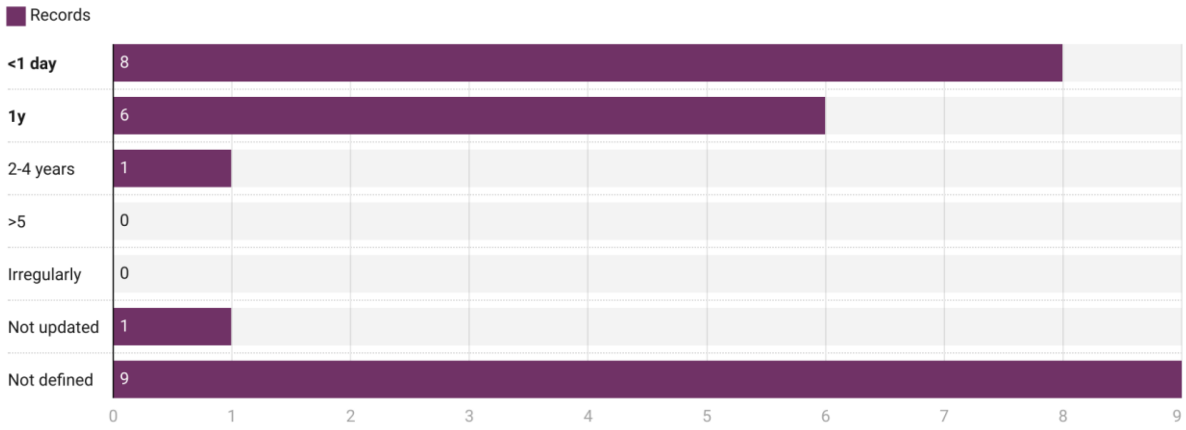



Figure 40: Distribution of the number of data sources per the user-determined classes of update frequency

# 6 Play 4 – Access the policy, regulatory, and ethical context

## 6.1 Background

Assessing the ethical considerations associated with the data sources is critical. In the following chapters, we examine the degree of openness and associated license schemes for the identified data assets. This examination also provides insights into how FAIR (Findable, Accessible, Interoperable, Re-usable) the datasets are. Regarding openness, the classification of the data spectrum as used by ODI will be employed:[56]

- **Open Data:** data that anyone can use, for any purpose, for free
- **Shared Data:** data that has pre-emptive licences for specific use cases, and open data descriptions of both the data and the licence conditions
- **Closed Data:** data that requires a per-use, custom licence, negotiated on a case-by-case basis (or cannot be shared or licenced).

## 6.2 Summarizing the Ethical Content

### 6.2.1 Data openness in Urban ReLeaf cities

In general, there is a tendency for European cities to adopt open-access practice and reduce the effect of data silos. We are witnessing a proliferation of open access ecosystems, with cities advocating for data sharing to maximize impact and value. Major efforts have been spent through several organisations, such as the ODI and the OKF[57] to bring up a concrete definition of "Open Data". The World Bank[58] has elaborated further on this subject, by including two additional dimensions regarding data openness:

- The data must be **legally open**, denoting the definition of the legal content before democratizing the data asset in the public domain.
- The data must be **technically open**, referring to the dissemination of data in electronic formats that are machine-readable and non-proprietary, so that anyone can access and use the data using common, freely available software tools. Data must also be publicly available and accessible on a public server, without password or firewall restrictions. To make Open Data easier to find, most organizations create and manage Open Data catalogues.

An important factor in open data is that it is published freely given the general public access to its content and to provide feedback. This factor is critical to the success of Urban ReLeaf that not only relies on the value of using such existing data, but also strives to reciprocate open publishing of data generated by the project.

In our analysis of the level of openness across the identified datasets we see promising results across Urban ReLeaf pilot cities (Figure 41).There is significant potential, as 80% and even more of the collected data assets are categorized to be fully **legally open** and democratized in the cities of Mannheim (80%), Cascais (82%), and Dundee (86%), using several technological tools, with the most common to be the data repositories (see chapter 5.2).



[56] https://www.theodi.org/about-the-odi/the-data-spectrum/
[57] https://icebreakerone.org/open-shared-closed/
[58] http://opendatatoolkit.worldbank.org/en/essentials.html

Utrecht reveals the most commitment to this goal with the entirety of its declared data sources to be openly accessed. Utrecht should be described as a frontrunner in the new vision of the Open Data "Governmental" Ecosystem (ODGE)[59], which has gained significant growth and acknowledgement.

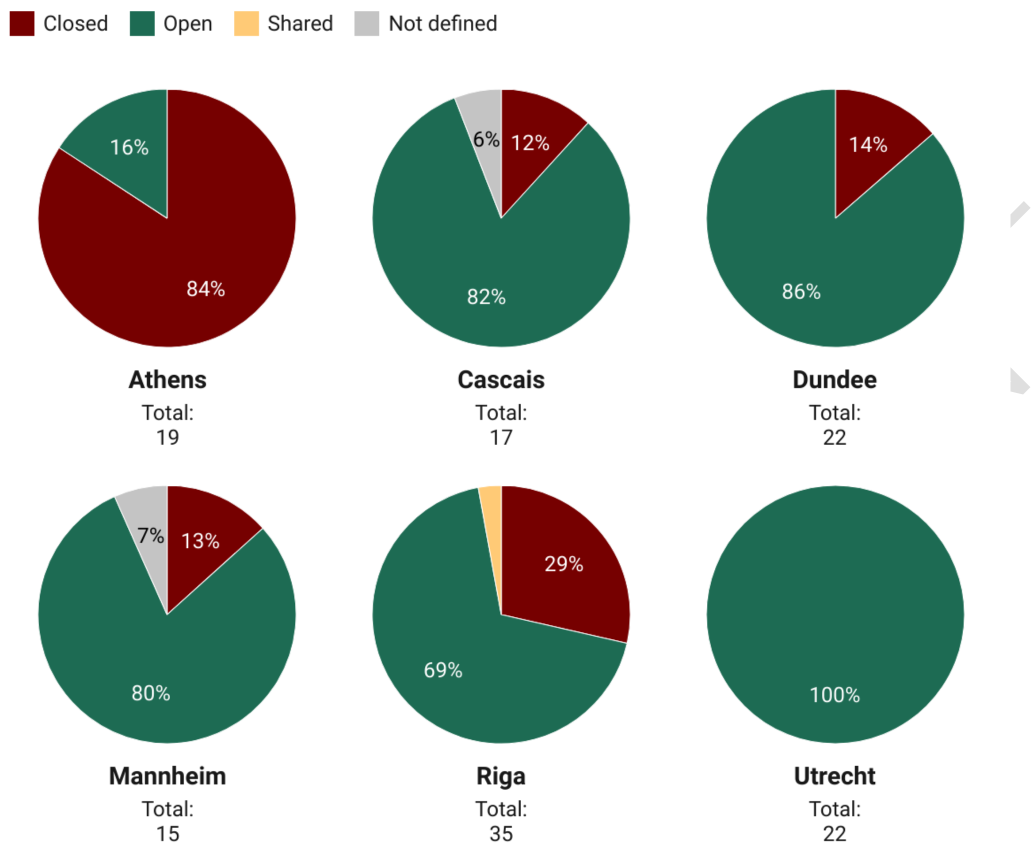



Figure 41: Distribution of the number of data records in the three classes of the data spectrum (open, shared, close).

A dedicated effort is given by Riga (69%), which showcases its tendency to disseminate its sources publicly. Nonetheless, we should indicate that since the number of datasets identified in Riga is greater compared to the other cities, this percentage corresponds to 24 actual sources. Finally, Athens is in the position of lower adoption of open assets, with almost 16% disseminated freely, compared to 84% of the datasets provided under a specific contract. Nevertheless, this outcome only expresses the data sources that are relevant to the project, and cities envisaged goals, thus we presumed that it does not express the whole picture of the city. Several public and private stakeholders across the country constantly join forces to provide curated data sources that aim to improve the quality of life of the societies and foster citizens' engagement. A typical example of such initiatives is the formulation of the PANACEA consortium.

Furthermore, reviewing the second aspect of openness, which is **technical openness**, it appears that the 6 cities have developed Open Data platforms and Catalogues in order to disseminate openly the data assets to the public. In Chapter 5.2, we reviewed the

[59] https://www.mdpi.com/1999-5903/4/4/900

technological options that each city has adopted for the accomplishment of such an objective. Table 16 briefly showcases these portals.

Table 16: Brief representation of the Open Data Portals and Catalogues that each reviewing city owns.

| City | Open Data Catalogues and Geoportals |
|---|---|
| Dundee | • Data Dundee city gov.uk (https://data.dundeecity.gov.uk/) |
| Riga | • Latvian's Open Data portal (https://data.gov.lv/eng)<br>• GEO RĪGA (https://georiga.lv/portal/apps/sites/#/georiga) |
| Athens | • Athens Geoportal (http://gis.cityofathens.gr/) |
| Mannheim | • StadtMannheim (https://mannheim.opendatasoft.com/page/home/) |
| Cascais | • Cascais City Hall (https://dadosabertos.cascais.pt/pt_PT/) |
| Utrecht | • Data portal of the municipality of Utrecht (https://data.utrecht.nl/)<br>• Geo Point of the province of Utrecht<br>• (https://geo-point.provincie-utrecht.nl/pages/open-data) |

All the examined open data catalogues have adopted the federated data cataloguing model[60], in which several decentralized ministries and departments of the city are responsible separately for the collection and evaluation of the stewarded data source, which later on is disseminated in the central portal. This statement can be proven by reviewing that several discrete organisations appear in each portal, (e.g. in Figure 42). It appears that the cities follow the best practices for the development of these catalogues, utilizing open-source solutions like CKAN[61] and Geonode[62], both of which address the most common features (Table 17) that ensure the curation and the interoperability of the information given.



60 http://opendatatoolkit.worldbank.org/en/technology.html#models
61 https://ckan.org/
62 https://geonode.org/

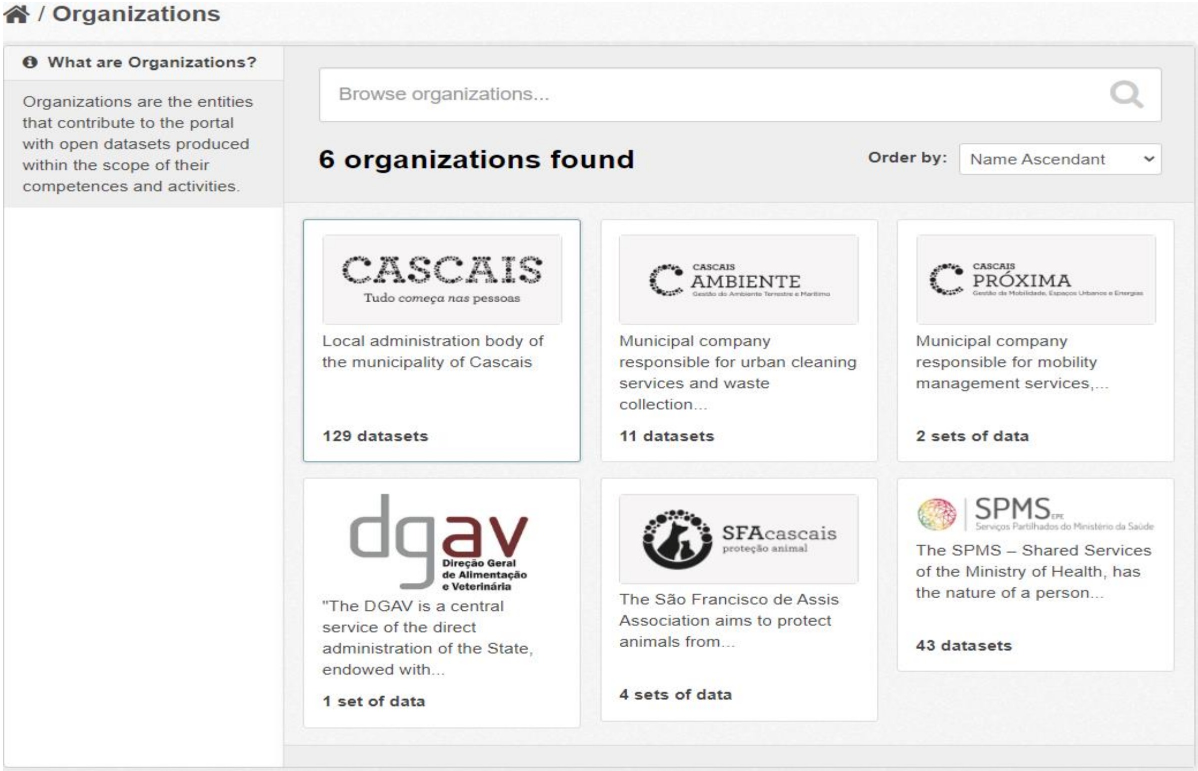



Figure 42: Indicative example of the decentralised open data cataloguing model that is adopted in the Cascais city, in which as it can be shown 6 main organisations are significantly contributing to the generation of the data sources.

Table 17: Common features that open data catalogues should have[63].

| Feature | Description |
|---|---|
| Easy access | Open Data catalogues make it very easy for users to access data quickly, freely and intuitively, without registration or login. |
| Search | Data catalogues should sort data by subject, organization or type, and support full-text searching of catalogue contents. |
| Machine-readable data access | Data are available for download in machine-readable, non-proprietary electronic formats. |
| Metadata | Key metadata, such as publication date and attribution, are prominently displayed for each dataset. |
| Clear data licences | Data licences are clearly and prominently displayed for each dataset, (see Chapter 6.2.2). |
| Data preview/visualization | Data catalogues should include some facility to preview the data prior to downloading or visualize the data using built-in graphing or mapping tools. |
| Standards compliance | Data catalogues have built-in support for various standards, such as data formats (e.g., CSV, XML, JSON) and metadata (i.e., Dublin Core). |
| API | APIs allow software developers to access the Open Data catalogues, and the data itself through software |
| Security | Data catalogues should implement security measures to protect data and metadata from being changed by unauthorized users. |

[63] http://opendatatoolkit.worldbank.org/en/technology.html#platforms

### 6.2.2 Open-accessed licences

In the second aspect of the data ethical canvas, it was essential to further under which licenses the data sources have been published in the open data portals. Addressing this aspect is important, for the technical partners of the Urban ReLeaf consortium to know beforehand the conditions under which these data sources may be used, and the legal content behind them, to avoid any unfortunate cases of data misuse. As it is illustrated in Figure 43 and further examined below, several standard and governmental licence schemes have been adopted by the countries.

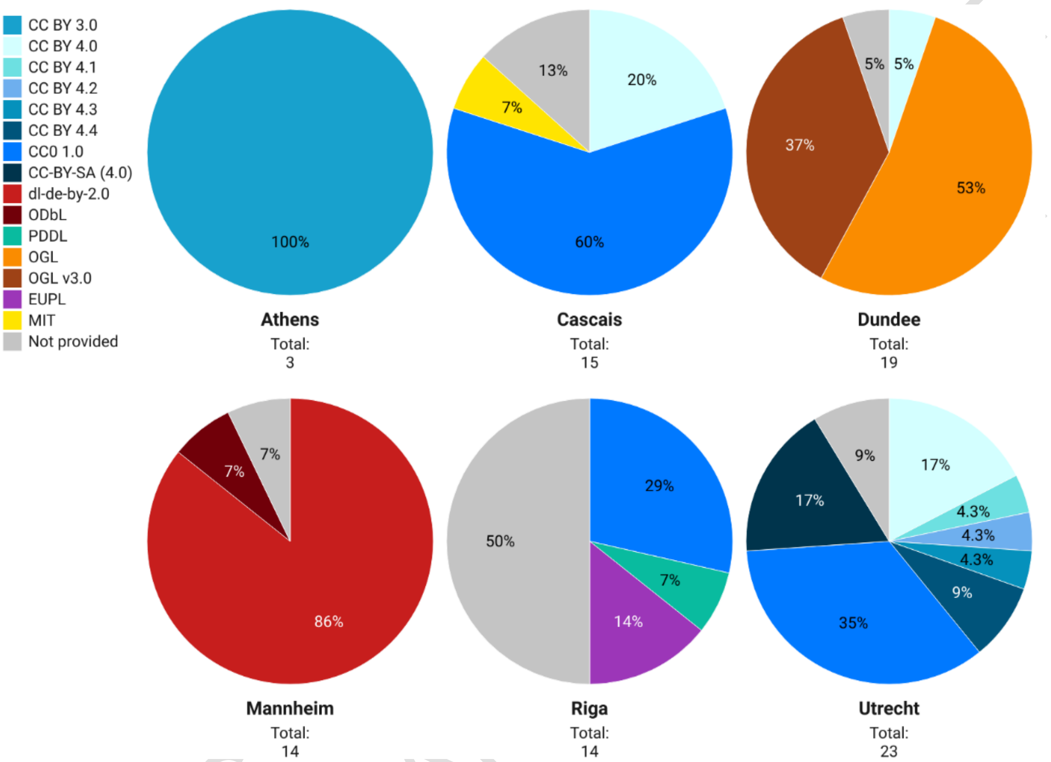



Figure 43: Open-access licences of identified datasets across pilot cities

For the first category, the Creative Commons (CC) and the Open Data Commons Open Database License (ODC-ODbL) have been chosen with different versions of their initially defined framework, e.g. CC0 1.0, CC BY (3.0-4.4), CC-BY-SA 4.0, and ODbL 1.0, (Table 18), with the most favourable being the CC0. A brief description of their content and differences are listed below.

Table 18: Distribution of the open-accessed licence under which data were published.

| Licences | Athens | Cascais | Dundee | Mannheim | Riga | Utrecht | Total |
|---|---|---|---|---|---|---|---|
| CC BY 3.0 | 3 | 0 | 0 | 0 | 0 | 0 | 3 |
| CC BY 4.0 | 0 | 3 | 1 | 0 | 0 | 4 | 8 |
| CC BY 4.1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 |
| CC BY 4.2 | 0 | 0 | 0 | 0 | 0 | 1 | 1 |
| CC BY 4.3 | 0 | 0 | 0 | 0 | 0 | 1 | 1 |
| CC BY 4.4 | 0 | 0 | 0 | 0 | 0 | 2 | 2 |
| CC0 1.0 | 0 | 9 | 0 | 0 | 4 | 8 | 21 |
| CC-BY-SA 4.0 | 0 | 0 | 0 | 0 | 0 | 4 | 4 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **ODC-ODbL 1.0** | 0 | 0 | 0 | 1 | 0 | 0 | 1 |
| **PDDL** | 0 | 0 | 0 | 0 | 1 | 0 | 1 |
| **MIT** | 0 | 1 | 0 | 0 | 0 | 0 | 1 |

Starting with the CC licences, in general, there are six different CC (i.e., CC BY, CC BY-SA, CC BY-NC, CC BY-ND, CC BY-NC-SA, CC BY-NC-ND)[64] that indicate the framework and the level of freedom of the wider public audience (i.e., citizens, private companies, or public entities) to data usage, transformation, maintenance, redistribution and commercial exploitation under potentially another licence. Figure 44 illustrates a high-level graphical description of the legal content that each CC licence represents.

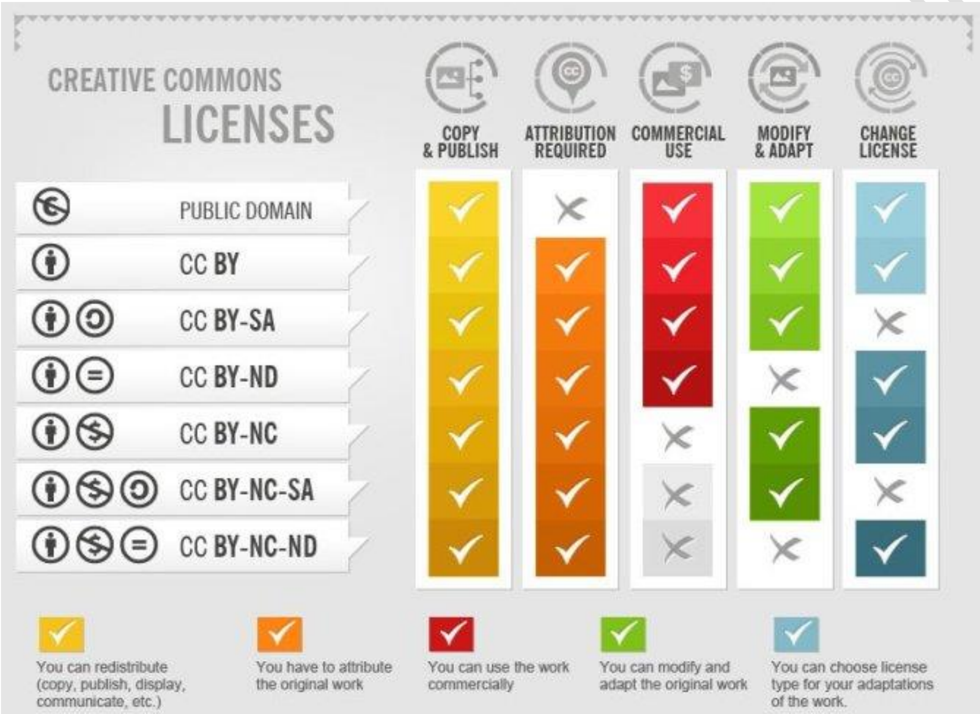



Figure 44: High-level graphical representation of the legal content behind the different CC licence schemes[30].

In terms of simplicity, the following bullet points will attempt to briefly present the legal content behind the open-accessed licences that appeared on the identified data sources, and there is a potential interest of the technical partners of the UR consortium to exploit them.

- CC0 is denoted as a **"No Rights Reserved"** licence that allows you to copy, modify, distribute and perform the work, even for commercial purposes, all without asking permission.
- CC-BY is denoted as an **"Attribution"** licence that allows distribution, remixing, adapting, and building upon the data owner's work, even commercially, as long as the users credit the data owner for the original creation.
- CC BY-SA is denoted as an **"Attribution-ShareAlike"** licence and allows to remix, adapt, and build upon data owner's work even for commercial purposes, as long as they credit the user and license their new creations under identical terms. This license is often compared to "copyleft" free and open-source software licenses. All new works

[64] https://foter.com/blog/how-to-attribute-creative-commons-photos/

based on the data owners will carry the same license, so any derivatives will also allow commercial use.

- CC BY-ND is denoted as an **"Attribution-NoDerivs"** licence that allows reusing the work for any purpose, including commercially; however, it cannot be shared with others in adapted form, and credit must be provided to the data owner.
- CC BY-NC is denoted as an **"Attribution-NonCommercial"** licence that allows remixing, adapting, and building upon the data owner's work non-commercially, and although their new works must also acknowledge the data owner and be non-commercial, they don't have to license their derivative works on the same terms.
- CC BY-NC-SA is denoted as an **"Attribution-NonCommercial-ShareAlike"** licence that allows remixing, adapting, and building upon the owner's work non-commercially, as long as they credit the data owner and license their new creations under identical terms.
- CC BY-NC-ND is denoted as an **"Attribution-NonCommercial-NoDerivs"** licence and it is the most restrictive of our six main licenses, only allowing others to download the data owner's works and share them with others as long as they credit the data owner, but the users can't change them in any way or use them commercially.

Subsequently, CC has also produced three open solutions specifically for data, datasets and databases, i.e. Open Data Commons Attribution Licence ODC-BY, (compatible with CC-BY), Open Data Commons Open Database Licence ODC-ODbL (compatible with CC BY SA), and Public Domain Dedication Licence PDDL (compatible with CC0)[65].

There is a single occasion, where the data source can be provided by performing a subsequent use of the openly accessed code. For the specific case of Cascais, and through the world data league (WDL) competition[66], two individual scientists developed a predictive model that can access the air quality index status of Cascais. The whole coding project is disseminated as a Jupyter Notebook python data format (.ipynb), and under the WDL Gitlab repository[67] following the Massachusetts Institute of Technology (MIT) licence scheme. Briefly described, by the MIT licence, the user can free-of-charge use, copy, modify, merge publish, distribute, sublicense and commercially exploit the source code or software, upon its mandatory inclusion of a copyright.

Furthermore, for the second category, governments and international organisations have developed bespoke licences for their data sources to gain greater recognition with free dissemination to the users, and therefore achieve a more interoperable concept with easy compliance. Some of these licences that have been employed by the pilot cities are presented in Table 19, with their brief descriptions presented in the following.

Table 19: Distribution of the open data assets according to the governance open-accessed licence under which they were published.

| Licences | Athens | Cascais | Dundee | Mannheim | Riga | Utrecht | Total |
|---|---|---|---|---|---|---|---|
| dl-de-by-2.0 | 0 | 0 | 0 | 11 | 0 | 0 | 11 |
| OGL | 0 | 0 | 10 | 0 | 0 | 0 | 10 |
| OGL v3.0 | 0 | 0 | 7 | 0 | 0 | 0 | 7 |

[65] http://www.discovery.ac.uk/files/pdf/Licensing_Open_Data_A_Practical_Guide.pdf
[66] https://www.worlddataleague.com/
[67] https://gitlab.com/worlddataleague/wdl-solutions

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **EUPL** | 0 | 0 | 0 | 0 | 2 | 0 | **2** |

- **OGL (v3.0)**: The National Archives has launched the United Kingdom Open Government Licence (OGL) facilitating the reuse of government and other public sector data sources. Under this specific licence, users are allowed to copy, publish, transmit and re-distribute the information, as well as to commercially exploit it. A specific copyright should be attributed, acknowledging the data owner. Certain exceptions appear for the exploitation of sensitive data.
- **dl-de-by-2.0**: The data licence of Germany is attributed with a common legal content as the CC-BY, allowing to reuse the work for any purpose, including commercially; however, ensuring that proper copyright attributed to the data owner will be included.

There are two cases in Riga city, where the attribution of the data source was provided as part of the web portal[68] owned by the statistical agency of Latvia, where the data sources can be freely downloaded in standard data formats. However, there is no reference for the licencing scheme that has been adopted for the data assets, whereas the licence that appears is the **European Union Public Licence (EUPL)** is solely related to the open-access usage of the software. An arbitrary assumption can be made for the accessing scheme of the data sources provided as the software itself is openly available, but certainly, an additional investigation is necessary. Subsequently, a similar investigation shall be applied in collaboration with the pilot cities for the following 13 cases (Table 20), which have been declared as open but a proper licence wasn't mentioned.

Table 20: Distribution per pilot city of the number of data sources that have been attributed as open-accessed with an absence of the open-accessed licence.

| Licences | Athens | Cascais | Dundee | Mannheim | Riga | Utrecht | Total |
|---|---|---|---|---|---|---|---|
| Not provided | 0 | 2 | 1 | 1 | 7 | 2 | **13** |

## 6.2.3 Trustworthy data sources

Data trustworthiness underpins the identification of steps and methodological procedures that an organisation should follow to safeguard the quality and curation of the collected information[69]. Assessing trust could entail two notions, which can be denoted as **"inner trust"** and **"outer trust"**. The first case indicates the stage when an organisation is about to launch a new service, for instance, and is essential to assess the trustworthiness of its data practices and document the results internally to provide peace of mind and confidence that it is doing everything possible to be a trustworthy steward of data. Later, once that service is up and running that same organisation might want to assess the degree to which it is trusted by its partners and customers (**outer trust**). Such an initiative can be performed by the constant engagement of the stakeholders in assessment workshops to request their feedback and further considerations of improvement or can be further expanded to the wider public, where these viewpoints could be stated upon digital survey forms, with close (i.e., ratings) or open questions, (Figure 45).



[68] https://geo.stat.gov.lv/stage2/#
[69] https://www.theodi.org/article/introducing-the-odi-trustworthy-data-stewardship-guidebook/

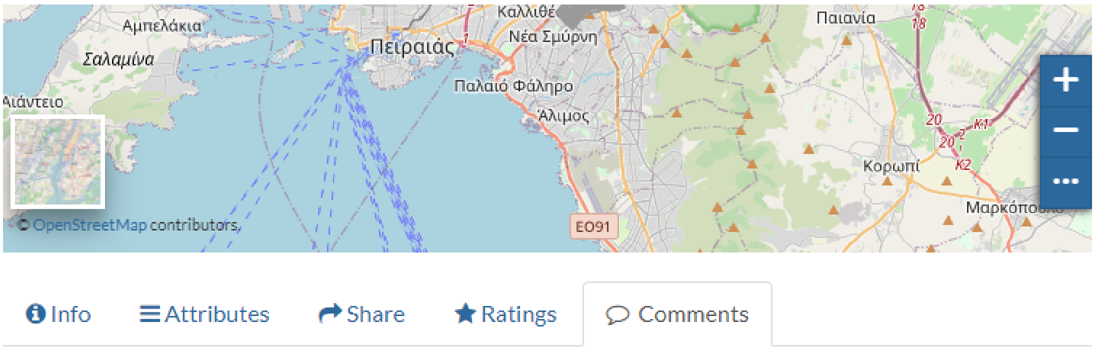



Figure 45: An indicative example of the Athens GeoNode open access data portal, which allows the wider public to rate the data asset and express any comment.

Apart from this general approach of requesting and more importantly giving the freewill of receiving a critical examination of the level of quality, there are general quality assessment standards that could warranty the data and metadata curation, and these concern aspects related to (i) the accuracy and reliability, (ii) timeliness and punctuality, (iii) accessibility and clarity, and (iv) coherence[70]. Under a similar dimension, in 2016, the FAIR guiding principles for scientific data management and stewardship were launched[71], as upon the successful completion of the above-mentioned goals, each data asset could be characterised as Fairness, Accessible, Interoperable and Reused.

We identified several occasions where the data steward provides details upon the generation of the produced asset that has been disseminated or expressed any reluctance. The lattert usually appeared in cases, where the data steward was not responsible for the generation of the data source, a rather common condition in every city with a decentralised open data cataloguing system.

Nevertheless, in Athens , two datasets existed, oriented on the air quality monitoring, which have been certified that, prior to their public dissemination, the observations are calibrated and certified by the EN ISO 17025[72]. For the same research domain, in Cascais, the outcomes of the air quality models generated during a hackathon contest have been evaluated using statistical validation metrics and thus strengthening their statistical importance and thus their quality. In Mannheim, a specific data quality assessment is applied for 6 data assets, which therefore is not publicly disseminated, thus further analysis is foreseen. In Dundee, biodiversity data sources have been produced following the specific directives of the EC in Birds and the Natura 2000. In addition, the Scottish air quality database have UKAS accredited Quality Assurance, Quality Control (QAQC) audits carried out every 6 months[73]. On the contrary to the aforementioned QAQC practices, three soil-related data assets are subjective to employed gathering procedures.

Furthermore, in Riga, 3 data sources related to air quality (e.g., the air quality measurements of in-situ stations) are declared to be produced according to the respective EU regulations; Directive 2008/50/EC on ambient air quality and cleaner air for Europe; 2011/850/EU: Commission Implementing Decision of 12 December 2011 laying down rules for Directives



[70] http://opendatatoolkit.worldbank.org/en/supply.html#general
[71] https://www.go-fair.org/fair-principles/
[72] https://ypen.gov.gr/perivallon/poiotita-tis-atmosfairas/dedomena-metriseon-atmosfairikis-rypansis/
[73] https://www.scottishairquality.scot/laqm/certificates-calibration

2004/107/EC and 2008/50/EC of the European Parliament and of the Council as regards the reciprocal[74]. Under the same rationale, the Freeport Authority of Riga is compliant with the ISO 14001 directive for environmental management so as to provide quality air quality observations[75]. Finally, the air quality zone map for NO2 and PM10, and the air quality index, ultraviolet (UV) radiation and pollen concentrations geospatial products have been produced utilising the ADMS-Urban[76], which is a comprehensive system for modelling air quality, and thus safeguard their accuracy.

Continuing with Utrecht, four consecutive citizen science datasets produced by the Snifferbike and Green Bicycle Roots projects and one provided by the city of Utrecht, have been listed, with a limited verification process and thus further examination is a prerequisite before further using them. However, the air quality observations produced by the RIVM organisation are declared to be collected following the Clean Air Agreement[77] and Air Quality Monitoring Information model and calibrated collectively on an annual basis. Subsequently, several models are utilised, e.g. the PROBE model[78], the Stress Test guidelines of the Wageningen University[79], the Water Vision Nature model[80], etc., and the official IPCC reports, in order to produce a variety of parameters, such as climate change maps, and geospatial products indicating the impact of climate change (e.g. heat maps, warm nights, risk drought stress, etc.).

Examining the accessibility level and in general the compliance to the FAIR principles, the EU metadata quality assurance assessment rating[81] reveals the following results for the open data portals of each city. Such an official assessment is not available for the rest of the data access interfaces and thus any conclusions by the authors shall be avoided.



Table 21: EU metadata quality assurance assessment rating levels for each open data portal of the examined cites (at the country level).

| City | EU metadata quality assurance assessment rates | | | | | |
|---|---|---|---|---|---|---|
| | Findability | Accessibility | Interoperability | Reusability | Contextuality | Rating |
| Athens | 23 / 100 | 41 / 100 | 30 / 110 | 41 / 75 | 0 / 20 | Sufficient |
| Cascais | 30 / 100 | 77 / 100 | 20 / 110 | 40 / 75 | 15 / 20 | Sufficient |
| Riga | 62 / 100 | 25 / 100 | 12 / 110 | 14 / 75 | 7 / 20 | Bad |
| Open data portal | 28 / 100 | 49 / 100 | 69 / 110 | 30 / 75 | 4 / 20 | Sufficient |
| Mannheim | 78 / 100 | 36 / 100 | 19 / 110 | 24 / 75 | 5 / 20 | Sufficient |

74 https://www.eionet.europa.eu/aqportal/doc/IPR%20guidance_2.0.1_final.pdf
75 https://likumi.lv/ta/en/en/id/200712-regulations-regarding-ambient-air-quality
76 http://www.cerc.co.uk/environmental-software/ADMS-Urban-model.html
77 https://www.cimlk.nl/
78 https://www.klimaateffectatlas.nl/en/drought-stress
79 https://www.klimaateffectatlas.nl/en/perceived-temperature-heat-map
80 https://www.klimaateffectatlas.nl/en/drought-susceptibility-nature
81 https://data.europa.eu/mqa?locale=en

| | | | | | | |
|---|---|---|---|---|---|---|
| Utrecht | 81 / 100 | 6 / 100 | 3 / 110 | 8 / 75 | 4 / 20 | Bad |
| Dundee | 23 / 100 | 15 / 100 | 25 / 110 | 24 / 75 | 0 / 20 | Bad |

# 7 Contribution of the Open-Access Data Sources

## 7.1 Searching and Selection Strategy

In this final chapter, we pinpoint some additional data sources that have been identified through the GEO and other EU initiatives, or in general, over the Web that could potentially support the project's objectives. We formulated a well-structured rapid reviewing process, using key terms of interest (KToIs). In general, KToIs are a key setup during literature reviews as they synthesise the Boolean operators under which the search and select process is performed. So far and according to our best knowledge, there is not any official digital library or catalogue that is able to provide a thorough view of all the datasets that have been created and appear over the Web. Some key efforts have been made during the past few years though, by launching the open-accessed Zenodo community. However, it does not give yet a holistic view. Hence, for this exercise, we finally chose to leverage on the most known searching digital platform, Google.

To identify the valuable research KToIs, we thoroughly examined the content gained during the implementation of the first play of the DLP, in order to find commonalities between the cities' scenarios and the objectives that were identified during the proposal phase. The outcome of this process was to conclude with the following group of terms (Table 22), and thus proceed with the literature review.

Table 22: Group of the selected keywords used in our searching and selection strategy. The operator "AND" was used to combine the static with each of the fluctuating categories. The * is a truncation command for searching for the root of a word and then retrieve any alternate endings.

| Changing Static | Question Components | Search Terms |
|---|---|---|
| Fluctuating | Tree registry | "Tree"* OR "Canopy" OR "plant"* OR "invasive species" OR "flora species" |
| | Greening Layers | "Land cover"* OR "Vegetation cover"* OR "Forest cover"* or "green layer"* |
| | Air Quality | "Air Quality"* OR "Air pollutant"* OR "PM"* |
| | Bioclimatic thermal | "Temperature" OR "Humidity"* |
| Static | Sensing devices | "in situ" OR "Earth Observation" OR "Remote Sensing" OR "satellite" OR "citizen science" OR "crowdsourc"* |

Before reviewing the outcomes of this process, we should mention that normally all the terms that are selected usually remain static and all of them are utilised during the searching Boolean operation. They could be described as an ideal condition, only when a single subject of research is investigated. On the contrary, this strategy was not ideal in our case, and hence, a subsequent division was taken up to introduce two groups of keywords, the "Fluctuating" and the “Static”. Starting with the Fluctuating KToIs, this group contains question components that are separated following the four main use cases (see Chapter **Error! Reference source not found.**), and thus are going to be changed based on the use case we were performing each time in the reviewing process. However, the “Static” category solely encapsulates words that remained unchanged and thus utilised in every operation. As shown the latter include KToIs related to a sensing device that is responsible to collect the data.

In order to examine the validity of each initially identified data source and denoted it as “Relevant” or “Non-Relevant”, post-hoc Inclusion and Exclusion (In/Ex) criteria were formulated and applied for every single data record. The In/Ex criteria followed the key objectives of the DLP steps, i.e. to be classified as open or shared upon a prior registration for instance, to be easily accessible under standardised data formats and user-friendly or machine-friendly interfaces, and to be legitimate for the consortium or the wider public to be retrieved. To ensure a high level of impartiality, the criteria were defined at the beginning and prior to the selection process and remained unchanged during the whole process. The outcomes are presented in the section below.

## 7.2 Identification of open-accessed Data sources that Fit4Purpose

The outcomes of the Searching and Selection strategy were able to identify 13 data sources overall and subsequently 7 relevant to the tree registry use case, 6 for the air quality and temperature/humidity use cases. The following tables (Table 23 and Table 24) provide an overview of the context behind these data assets. The descriptions of the identified datasets are given based on the same categorisations and nomenclatures that were determined and applied for the quantitative analysis of the data ecosystem mapping of the cities’ data assets, (see Chapter 5.1). Hence, key aspects related to date, repeatability and the data spectrum classification will follow an identical harmonisation process.

Complementary to these, we were able to identify a scoping systematic literature review[82], in which a thorough investigation was implemented in various EO-based data assets to either produce the same classification product but in higher quality or to be exploited as a valuable feature during the training of ML models. The aspects that these datasets have covered are mostly related to the Land or Vegetation (i.e., Tree, Forest, and Agriculture) cover mapping. Subsequently, additional VHR benchmarking datasets are also presented and therefore declared as essential datasets with which the CNN model could be employed to cover the objectives.



Table 23: Open-accessed data sources that could be declared as a potential of interest for further exploitation in the context of the Urban ReLeaf project. These datasets are mostly related to “Use case 1: Participatory Tree Registry (including as well as the VHR) green layer that will be provided by the satellite EO data”.

| Tree-Registry and VHR greening layer Relevant Data Sources | | | |
|---|---|---|---|
| **Data source #1** | | | |
| **Source name** | iNaturalist | **Data owner** | iNaturalist |
| **Brief description** | iNaturalist is a community science platform, where participants record observations representing encounters with individual organisms. Each observation can have one or more photos as evidence of the encounter. | | |

[82] https://www.mdpi.com/2072-4292/14/5/1263

| | | | |
|---|---|---|---|
| | All observations are associated with a single user who recorded the observation. Much of the activity on the platform relates to identifying the single taxon that the organism represents. | | |
| **Data source to be explored** | Fauna and Flora species over the whole EU (including the UK) | | |
| **Date** | Not Defined | **Repeatability** | <1y, but maybe without covering the same locations |
| **Guidelines for data access** | iNaturalist provides open-access one of the world's largest public datasets of photos of living organisms, containing over 70 million photos. It is structured as a "bucket" of images, stored in AWS S3, in CSV data formats. The contents of the dataset include information of observations, observers, photos, and taxa. **GitHub project:** https://github.com/inaturalist/inaturalist-open-data | | |
| **Data Spectrum** | Open | **Open- Licensing** | CC0 1.0, CC BY 4.0, CC BY-NC 4.0 |
| **Data source #2** | | | |
| **Source name** | NBN Gateway | **Data owner** | National Biodiversity Network |
| **Brief description** | The National Biodiversity Network (NBN) Gateway is a "data warehouse" for biodiversity information, about particular species in the UK. Individual records, covering plants, mammals, birds and invertebrates, are stored on the NBN Gateway. | | |
| **Data source to be explored** | Fauna and Flora species repositories only in the UK | | |
| **Date** | Not Defined | **Repeatability** | 1y |
| **Guidelines for data access** | NBN atlas can be retrieved using a fully documented API (https://api.nbnatlas.org/). The data records are provided either in GeoJSON or CSV, depending on the data source. | | |
| **Data Spectrum** | Open | **Licensing Copyright** | CC0 1.0, CC BY 4.0, CC BY-NC 4.0 |
| **Data source #3** | | | |
| **Source name** | **EFISCEN inventory** | **Data owner** | **European Forest Institute** |
| **Brief description** | The European Forest Information SCENario model (EFISCEN) contains a forest inventory for 32 European countries: Albania, Austria, Belarus, Belgium, Bosnia-Herzegovina, Bulgaria, Croatia, Czech Republic, Denmark, Estonia, Finland, France, Germany, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Macedonia, Netherlands, Norway, Poland, Portugal, Romania, Russian Federation, Slovakia, Slovenia, Sweden, Switzerland, Turkey, and United Kingdom. | | |
| **Data source to be explored** | Forest types can be distinguished according to region, owner class, structure, site class and tree species. Forest type and age class, the forest area, the total and mean volume, the total annual increment and the current annual increment can be retrieved. | | |
| **Date** | Not Defined | **Repeatability** | Irregularly |
| **Guidelines for data access** | Upon registration data can be retrieved in CSV data format through the DB (http://dataservices.efi.int/efiscen/index.php) | | |

| | | | |
|---|---|---|---|
| **Data Spectrum** | Shared | **Licensing Copyright** | Reference to EFISCEN and the original data owner. Data access upon prion registration |
| **Data source #4** | | | |
| **Source name** | FISE | **Data owner** | FISE and EEA |
| **Brief description** | The Forest Information System for Europe (FISE) Data Catalogue delivers under the Spatial Data Infrastructure (SDI) infrastructure, official forest datasets from the statistical platforms of all the European countries. | | |
| **Data source to be explored** | The provided data assets are the following:<br>• **Tree cover density (10-100m):** High Resolution Layer: Tree Cover Density 2018 (raster 10m), Sep. 2020; High Resolution Layer: Tree Cover Density 2012 (raster 100m), Mar. 2018; High Resolution Layer: Tree Cover Density 2015 (raster 20m), Mar. 2018; High Resolution Layer: Tree Cover Density 2015 (raster 100m), Mar. 2018<br>• **Forest area:** Forest Area 2015 based on Copernicus HRL Forest products - version 1, Oct. 2018; Forest Area 2012 based on Copernicus HRL Forest products - version 1, Oct. 2018<br>• **Tree cover change mask:** High Resolution Layer: Tree Cover Change Mask 2015-2018 (raster 20m), Dec. 2020; High Resolution Layer: Tree Cover Change Mask 2012-2015 (raster 20m), Dec. 2020<br>• **Dominant Leaf Type (10-20m):** High Resolution Layer: Dominant Leaf Type 2018 (raster 10m), Sep. 2020; High Resolution Layer: Dominant Leaf Type 2015 (raster 20m), Apr. 2018; High Resolution Layer: Dominant Leaf Type 2012 (raster 20m), Apr. 2018<br>• **Dominant Leaf Type change (20m):** High Resolution Layer: Dominant Leaf Type Change 2015-2018 (raster 20m), Dec. 2020<br>• **Small woody features (raster: 5m or vector):** High Resolution Layer: Small Woody Features 2015 (raster 5m), Nov. 2019; High Resolution Layer: Small Woody Features 2015 (vector), Nov. 2019; High Resolution Layer: Small Woody Features 2015 (raster 100m), Nov. 2019 | | |
| **Date** | 2012-2020 | **Repeatability** | 2-4 y |
| **Guidelines for data access** | Six different types of data formats are utilised in order to deliver different information on the aforementioned content, PDF format under which the documentation and the statistical report are stored and provided, CSV format including the tabular statistical data, and for the spatial datasets, several options are available, GeoTIFF or ESRI shapefiles, as well as machine-friendly OGC-compliant web services (WMS, WCS, and WFS). Additionally, a RESTful API has been deployed and fully documented (https://sdi.eea.europa.eu/catalogue/doc/api/index.html) | | |
| **Data Spectrum** | Open | **Licensing Copyright** | Reference to the data owner and EEA is required |
| **Data source #5** | | | |
| **Source name** | Plant Image Search | **Data owner** | Arnold Arboretum of Harvard University |
| **Brief description** | The Arnold Arboretum plant image (ArbPIX) database is a free resource of historical and contemporary images of accessioned plants. Images detail structural characteristics (morphology) and seasonal aspects (phenology) of temperate woody plants in the living collections. The data collections are provided under open and reasonable access, predominately for conservation, education, and horticultural display purposes (source). | | |

| | | | |
|---|---|---|---|
| **Data source to be explored** | The data repository contains georeferenced high-resolution photos of plant species and their taxa classification label, country of their location | | |
| **Date** | 150 years (up to 2022) | **Repeatability** | Not Defined |
| **Guidelines for data access** | Provided either through OGC-compliant services or as Excel spreadsheets (https://arboretum.harvard.edu/research/data-resources/). | | |
| **Data Spectrum** | Shared | **Licensing Copyright** | CC BY-NC 4.0 |
| **Data source #5** | | | |
| **Source name** | GROW Observatory EPD | **Data owner** | IIASA |
| **Brief description** | The Edible Plant Database (EPD) is an outcome of the GROW Observatory, a European Citizen Science project on growing food, soil moisture sensing and land monitoring. This dataset was disseminated via a web service hosted by IIASA (Austria) that provides information about suitable plants for locations within Europe, based on climate. The web service was used in the GROW app, this dataset is the base information used for that service. | | |
| **Data source to be exploited** | 15 growing parameters with a supporting database on germination of 146 edible plant species included in EPD. Planting calendars and germination data to all 12 European climate zones are also included with high-resolution images purchased under licence. | | |
| **Date** | 2020 | **Repeatability** | Unspecified |
| **Guidelines for data access** | Data can be retrieved in Excel data format from the open data repository (https://discovery.dundee.ac.uk/en/datasets/edible-plant-database) | | |
| **Data Spectrum** | Open | **Licensing Copyright** | CC-BY 4.0 |
| **Data source #6** | | | |
| **Source name** | ForestGEO | **Data owner** | Smithsonian Institution |
| **Brief description** | The Forest Global Earth Observatory (ForestGEO) is a global network of scientists and forest research sites dedicated to advancing the long-term study of the world's forests. Their distribution covers the whole globe with already defined sites over some European countries that are of interest, such as the Netherlands, Germany and the UK. | | |
| **Data source to be exploited** | The tree data represents information of all free-standing trees with a DBH of at least 1 cm, which are tagged, measured, and identified to species. Additionally, key variables related to habitat and the association with tree species and topography (i.e. convexity, slope and mean elevation) of where each data record has been identified, are also available along with photos illustrating their natural representation. | | |
| **Date** | Not Defined | **Repeatability** | >5y |
| **Guidelines for data access** | Data can be requested via the online data repository of ForestGEO, where registered users have to fill out a data request form. Dedicated programming projects have been also developed and freely disseminated in the GitHub repository (https://github.com/forestgeo), formulating examples of accessing these data sources and analysing the different variables that are being collected. Furthermore, it is declared that in general data shall be used for research and knowledge purposes, | | |

| | | | |
|---|---|---|---|
| | prohibiting any commercial exploitation, and alteration or modification of the content based on the initial description given during the submission of the data request form. | | |
| **Data Spectrum** | Shared | **Licensing Copyright** | **Citation is necessary and is compliant with the terms of the user** |
| **Data source #7** | | | |
| **Source name** | **Data.Geo-Trees** | **Data owner** | GEO-Trees initiative |
| **Brief description** | DATA.GEO-TREES, formerly named Forest Observation System is an international cooperation, which envisages establishing a global in-situ forest biomass database in order to support EO and encourage investment in relevant field-based observations and science. | | |
| **Data source to be exploited** | DATA.GEO-TREES provides well-curated above-ground biomass (AGB) estimation of 260 forest data plots globally in a unified format that is aggregated from tree-level data consistently. Several parameters are collected, such as tree records (above a certain Diameter Breast Height (DBH)), along with their species, DBH and height, and accurate GPS location of each observation. | | |
| **Date** | 1980-2017 | **Repeatability** | Irregularly (varies in different countries) |
| **Guidelines for data access** | Data can be accessed from the open data repository of Data.Geo-Trees, in EXCEL data format. | | |
| **Data Spectrum** | Open | **Licensing Copyright** | CC-BY 4.0, pre-determined citation should be declared prior to its usage |

Table 24: Open-accessed data sources that could be declared as a potential of interest for further exploitation in the context of the UR project. These datasets are mostly related to *Use Case 3: Measurement of temperature and humidity and Use case 4: Measurement of air quality.*

| | | | |
|---|---|---|---|
| **Air Quality, Temperature and Humidity** | | | |
| **Data source #1** | | | |
| **Source name** | **Netatmo connect** | **Data owner** | **Netatmo** |
| **Brief description** | Netatmo is a network with smart indoor and outdoor sensors that give continuous observations of several air quality-related variables. The network expands all over the globe (https://weathermap.netatmo.com/). | | |
| **Data source to be exploited** | Smart air quality monitoring observations (i.e., temperature (ºC), humidity (%), CO2 (ppm) and Noise (dB)[83]). | | |
| **Date** | 2016-2023 | **Repeatability** | <1d |
| **Guidelines for data access** | A RESTful API is provided in order to retrieve those measurements, to which only authorized users can have access. A particular GitHub project (https://github.com/NINAnor/cityTairMapping) is available that provides a Python script for data access over the specified region. The observations | | |

[83] https://dev.netatmo.com/apidocumentation/aircare

| | | | |
|---|---|---|---|
| | are provided in CSV or XLS data format[84]. Maximum 500 observations per API request. | | |
| **Data Spectrum** | Shared | **Licensing Copyright** | Can be accessed for commercial use but prior authorization is required[85]. |
| **Data source #2** | | | |
| **Source name** | hackAIR | **Data owner** | HackAir consortium |
| **Brief description** | HackAIR EU H2020 project attempted to raise collective awareness about the daily conditions of air quality and thermal comfort, as well as provide information about the probability of forest fires in Europe. It enables you to easily access information relevant to outdoor air pollution and thermal comfort conditions and also to contribute to their monitoring by stating your perception about them. In addition, hackAIR informs you of the probability of a forest fire in your area. | | |
| **Data source to be exploited** | PM2.5 and PM10 measurements (μg/m$^3$) are available for the whole European continent and beyond. In the open-accessed repository (i.e., Zenodo) there is data that was collected by users using (i) low-cost sensors, as well as observation from stationary sensors, (ii) mobile HackAIR sensors, and (iii) air pollution estimations from photos (i.e., Flickr crowdsource photos, photos from webcams and sky photos that users upload on the hackAIR mobile application) depicting the sky. | | |
| **Date** | 2018 | **Repeatability** | Not updated |
| **Guidelines for data access** | HackAIR data can be exploited through the open-accessed repository of Zenodo (https://zenodo.org/record/2222342#.Y_NHinbP1D8), where the data has been uploaded as Excel files, containing the observation, the date/time is in UTC timezone and the point-wise coordinates. | | |
| **Data Spectrum** | Open | **Licensing Copyright** | CC0 4.0 |
| **Data source #3** | | | |
| **Source name** | EEA Air Quality | **Data owner** | EEA |
| **Brief description** | The European Environmental Agency (EEA) contains holds in-situ hourly observations of the main pollutants that are responsible for the air quality degradation, in each country. In EEA's data repository, we can find both E1a and E2a, with the first to be denoted as the validated data that are reported by each EEA member state every September for the whole year, and the second referring to the up-to-date data on hourly basis from most of its member states. | | |
| **Data source to be exploited** | Varied according to each member states air quality sensors' network. However, in every country, we can find measurements of PM 2.5 and 10. | | |
| **Date** | 2013-2023 | **Repeatability** | <1d |
| **Guidelines for data access** | The data is freely downloadable under CSV files via the EEA's developed platform (https://discomap.eea.europa.eu/map/fme/AirQualityExport.htm), where the user fills out a specific form in order to generate the URL. A more updated page with the developed service and a Python script can be | | |

[84] https://github.com/openfirmware/netatmo-dl
[85] https://dev.netatmo.com/legal

| | | | |
|---|---|---|---|
| | found also in https://discomap.eea.europa.eu/map/fme/AirQualityUTDExport.htm. | | |
| **Data Spectrum** | Open | **Licensing Copyright** | CC BY 4.0 |
| **Data source #4** | | | |
| **Source name** | European Industrial Emissions Portal | **Data owner** | EEA |
| **Brief description** | The Industrial Emissions Portal covers over 60,000 industrial sites from 65 economic activities across Europe. It is provided with the sites' location and administrative data, along with their releases and transfers of regulated substances to air, water and land, and transfers of waste. For large combustion plants (LCPs), there is more detailed data on energy input and emissions. | | |
| **Data source to be exploited** | Particulate Matter (PM) 2.5 and 10, as well as pollutant thresholds for releases per year (50.000 kg/y). | | |
| **Date** | 2007-2020 | **Repeatability** | 1y |
| **Guidelines for data access** | Raw data can be downloaded in CSV data format from EEA portals (https://industry.eea.europa.eu/explore/explore-data-by-pollutant or https://industry.eea.europa.eu/download). | | |
| **Data Spectrum** | Open | **Licensing Copyright** | Copyright of EEA |
| **Data source #5** | | | |
| **Source name** | CitiAIR | **Data owner** | CityMeasure consortium |
| **Brief description** | CitiAIR is an initiative with the CityMeasure EU-funded project, which attempted to create a different data repository, where different CS-related projects and stories on air quality are stored and disseminated. So far, 46 initiatives have been launched from 45 organisations, and 39462 sensing devices. The countries that have been declared so far of the potential to maintain relevant data sources are Greece, Netherlands, Germany, Portugal and UK, while concerning the organisations claimed of stewarding the data sources, are city stakeholders, universities and research institutes, companies in the private sector, and governmental and non-governmental organisations. A closing remark is to mention that the projects that appeared in the dashboard are categorised as "ongoing" or "inactive/ended", and this information arbitrarily provides a preliminary insight into how updated are the collected data assets. | | |
| **Data source to be exploited** | Data that could be explored are in general the user CS projects that have been implemented, technologies that have been adopted and best practices upon the citizens' engagement aspect. By exploiting this inventory, additional EU projects can be identified and thus parameters that are measured (e.g. air pollutants), and geospatial AQ data sources and access interfaces. The AQ variables that have been identified to be covered so far by the projects are Black Carbon, CO2, Humidity, NO2, Noise, O3, PM1, 2.5 and 10, Temperature, Ultrafine particles, and VOCs. | | |
| **Date** | Not Defined | **Repeatability** | Not Defined |
| **Guidelines for data access** | An open accessed repository has been deployed, cataloguing the aforementioned information related to the AQ parameters that have been measured, the data owners, and the countries. Complementary to this, specific details are given regarding the existing projects under which the | | |

| | | | |
|---|---|---|---|
| | observations were collected and the technological interfaces that should be accessed in order to retrieve the data. This might be from platforms to GitHub projects. From the quick research that we did, the majority of the projects store their observations following standardised data formats (e.g. CSV or GeoJSON), while on some occasions API interfaces are given. | | |
| **Data Spectrum** | Varies based on each data provider adopted policy | **Licensing Copyright** | Varies based on each data provider adopted policy. |
| **Data source #6** | | | |
| **Source name** | Purple Air | **Data owner** | PurpleAir |
| **Brief description** | Purple Air provides low-cost and highly reliable PM2.5 sensors that can be deployed by anyone and collect observations in real time. Therefore, an open-accessed API and a web map application are provided, by which any user can have access and download the required measurements. The Purple Air Map uses the Air Quality Index (AQI) thresholds created by the EPA for their map interface, whereas sensors have been validated and certified by AC Sensor Performance Evaluation Center. | | |
| **Data source to be exploited** | The parameters that are provided are the humidity, temperature, air pressure, and PM1, 2.5, and 10. All the aforementioned can be retrieved both as real-time observations and in aggregated form. | | |
| **Date** | 1990-2023 | **Repeatability** | <1d |
| **Guidelines for data access** | Users can explore two access interfaces in order to retrieve the AQ observations, the Web map platform and the RESTful API, in which the user selects the specific sensor from which the data are desired to be downloaded and then perform either a download or a GET request action, so as to receive the corresponding CSV file. However, the API can be assessed upon a pricing contract. | | |
| **Data Spectrum** | Open (Web map only)<br>Closed (API) | **Licensing Copyright** | Not provided |

# 8 Conclusion

In this deliverable, we attempted to identify the existing data sources that appear within the six pilot cities of the Urban ReLeaf project, and the data owners that steward, manage and maintain these data assets. Addressing this aspect, the Data Landscape Playbook methodology was adopted, which comprised of four main steps that aim to map the data network of each city, identify soft and formal data flow exchanges, and highlight details related to the infrastructure, ethical content and level of openness. With these perspectives, the most known and standardised data formats are usually adopted, which, in most occasions are provided within data platforms. However, the most lightweight formats (geopackage) that are favoured by the SW engineering communities, or those that are often adopted upon a parallel computing process (e.g. Apache parquet[86]), seem to be left still uncovered.

Similar viewpoints also occur in the data interfaces, since the most "machine-useable" technologies and software (e.g. APIs) were not preferable in most cases. Furthermore, a rather positive result that was gained from this review, was the general tendency of the cities to provide the data assets openly. Standard open-accessed licences from Creative Commons appeared to be applied for most of the data assets in order to safeguard and define the content for their usage. Nevertheless, in rare occasions, this sense of openness is also accompanied by the subsequent openness in methods that have been applied to evaluate and further certify data accuracy. The latter seems to be an aspect of significant importance, as it has been described as a key factor to build trust over data and long-lasting engagement. As a final remark, it is worth advocating that the Data Ecosystem Mapping is an essential step in the process of finding collaborative pathways for citizen-powered observations to complement official data streams for holistic environmental monitoring.



[86] https://parquet.apache.org/

## References

Alexopoulos, C., Zuiderwijk, A., Charapabidis, Y., Loukis, E., & Janssen, M. (2014). Designing a second generation of open data platforms: Integrating open data and social media. In International Conference on Electronic Government (pp. 230-241). Springer, Berlin, Heidelberg.

Angelo, H. Added value? Denaturalizing the “good” of urban greening. Geography Compass. 2019; 13:e12459. https://doi.org/10.1111/gec3.12459Benz, S. A., & Burney, J. A. (2021). Widespread race and class disparities in surface urban heat extremes across the United States. Earth's Future, 9, e2021EF002016. https://doi.org/10.1029/2021EF002016

Charalabidis, Y., Alexopoulos, C., Diamantopoulou, V., & Androutsopoulou, A. (2016). An open data and open services repository for supporting citizen-driven application development for governance. In 2016 49th Hawaii International Conference on System Sciences (HICSS) (pp. 2596-2604). IEEE.

Harrison, T.M.; Pardo, T.A.; Cook, M. Creating Open Government Ecosystems: A Research and Development Agenda. Future Internet 2012, 4, 900-928. https://doi.org/10.3390/fi4040900

Karagiannopoulou, A., Tsertou, A., Tsimiklis, G. & Amditis, A. Data Fusion in Earth Observation and the Role of Citizen as a Sensor: A Scoping Review of Applications, Methods and Future Trends. Remote Sens. 14, (2022).

Kmail, Aseel & Onyango, Vincent. (2020). A GIS-based assessment of green space accessibility: case study of Dundee. Applied Geomatics. 12. 10.1007/s12518-020-00314

Petzold, J.; Mose, L. Urban Greening as a Response to Climate-Related Heat Risk: A Social–Geographical Review. Sustainability 2023, 15, 4996. https://doi.org/10.3390/su15064996

Richter, A. et al. Storytelling for narrative approaches in citizen science: towards a generalized model. J. Sci. Commun. 18, (2019).

Sethi, M., Lamb, W., Minx, J. & Creutzig, F. Climate change mitigation in cities: A systematic scoping of case studies. Environ. Res. Lett. 15, (2020).

United Nations Environment Programme and United Nations Human Settlements Programme (UN-Habitat) (2021). Global Environment for Cities-GEO for Cities: Towards Green and Just Cities. UNEP, Nairobi

# Appendix A: Data Ecosystem Maps

The following figures showcase the data ecosystem maps for each pilot city using the DLP methodology. Providing some additional details of its design, it mainly consists of two main aspects, (i) the **actors** (denoted with circles) and the **data assets** (denoted with squares). Analysing the first term, this group presents all the organisations that already steward one or more data assets. Subsequently, the data owners are divided into two sub-groups. The first sub-group are the city representatives and are denoted with blue coloured circles in the outer area of the circle and the second sub-group is the five technical partners of the UR consortium, who are located in the centre of DEMs and appeared with different colouring choices, based on which of the four use case categories they are involved in. The same colouring differentiation is adopted for the data assets. In the case a data asset could not be described by any of the four use cases, a grey colouring choice was adopted. Furthermore, a single-coloured rhombus shape was introduced in cases, where a coalition between several partners was identified (i.e., in terms of a funded project or a joined initiative).

Different connection lines have been introduced in these graphical representations, aiming to reveal different messages. Further details are given in the following bullet points. The primary goal behind those lines was to showcase on the one hand the "formal-exchanges" between the data assets and data owners that exist within each city, and on the other hand to perform some preliminary suggestions, the so-called "soft-exchanges", where the UR technical partners could contribute.

- **Light-Pink lines:** Denote the connections between the data owners of the city and the data assets.
- **Dark-Blue lines:** Denote the connections between the different sub-units, such as departments, agencies, governmental and non-governmental organisations that appear within a city and mainly are operated on behalf of the wider administration of each city (e.g. Dundee city council, Municipality of Athens, Province of Utrecht, etc.).
- **Orange lines:** Illustrate the identified connections between the data owners of the city, who collaborate in order to generate one or more datasets. This is an outcome that was retrieved during the reviewing process and the specific step of context screening in each data record. Nevertheless, we should declare that some deviations might appear upon this result, as the connections were defined solely by the authors of this document, without a cross-validation of this results from the cities.
- **Multicolour Dashed lines:** Follows the insights that are implied by the existing data assets and owners and tries to indicate some connections that could be formulated between the basic data actors of the city and the technical partners of the consortium. The aforementioned envisages highlighting some opportunities that could be explored. This exercise was sculptured by examining the local datasets and where the digital tools of UR would complement the work that is already done. An indicative example could be the hypothetical scenario, where IIASA and ICCS could provide data sources that give insights about the thermal status of an investigated area to the Architect's Office of Riga, which produces the UHI EO-based datasets to the Wageningen and Wittiveen+Bos organisations of Utrecht, which produce the Heat CC projection maps.

However, it should be noted we intentionally do not specify directions in these connections, since the project is still in its early stage.  We are not yet in a position to determine the exact data flows and make valid conclusions about the organisations involved. Such uncertainties led us to the final decision of solely illustrating the potential connections that we could explore in the long run of the project.

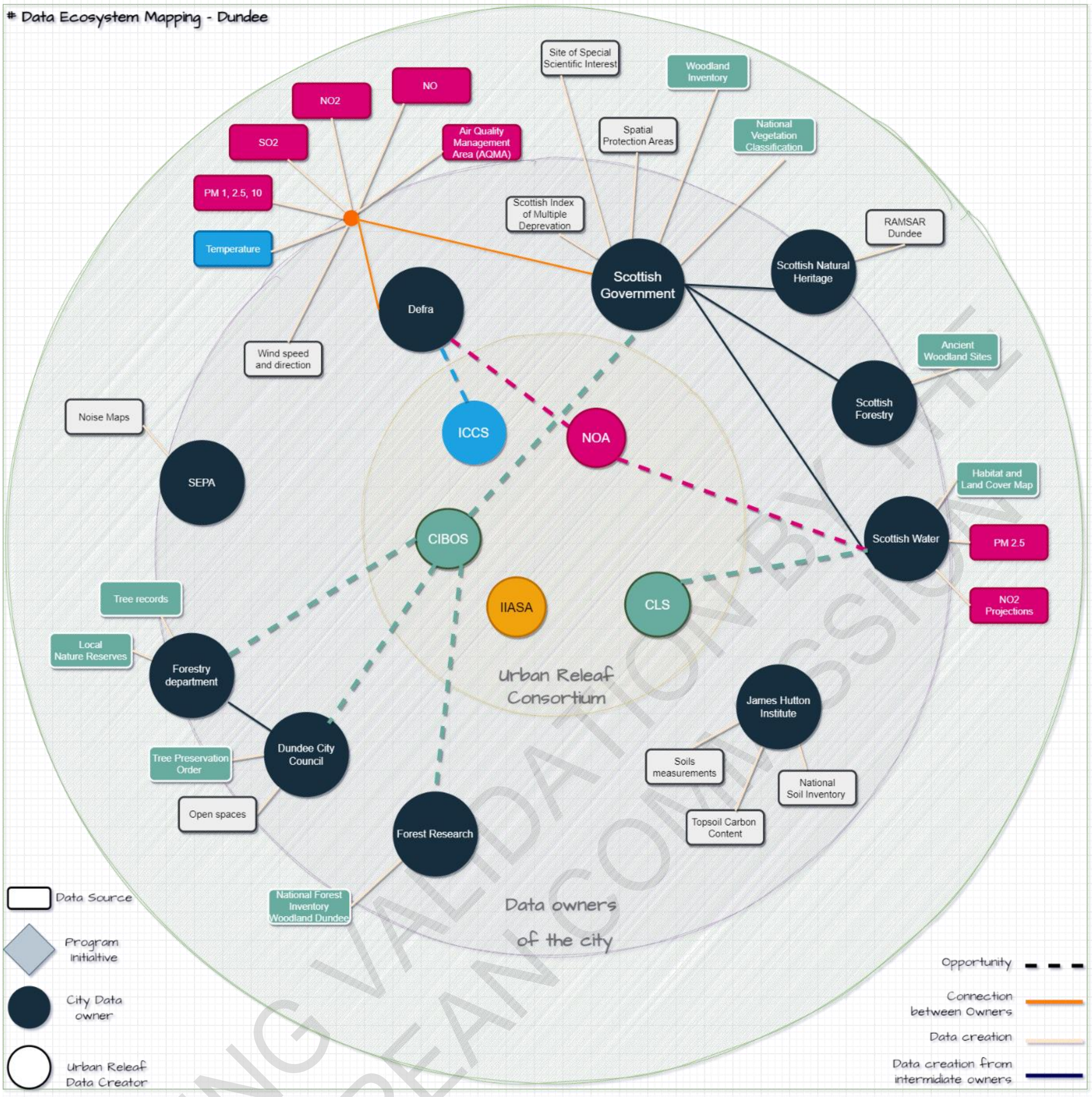



Figure 46: Data Ecosystem Map for Dundee

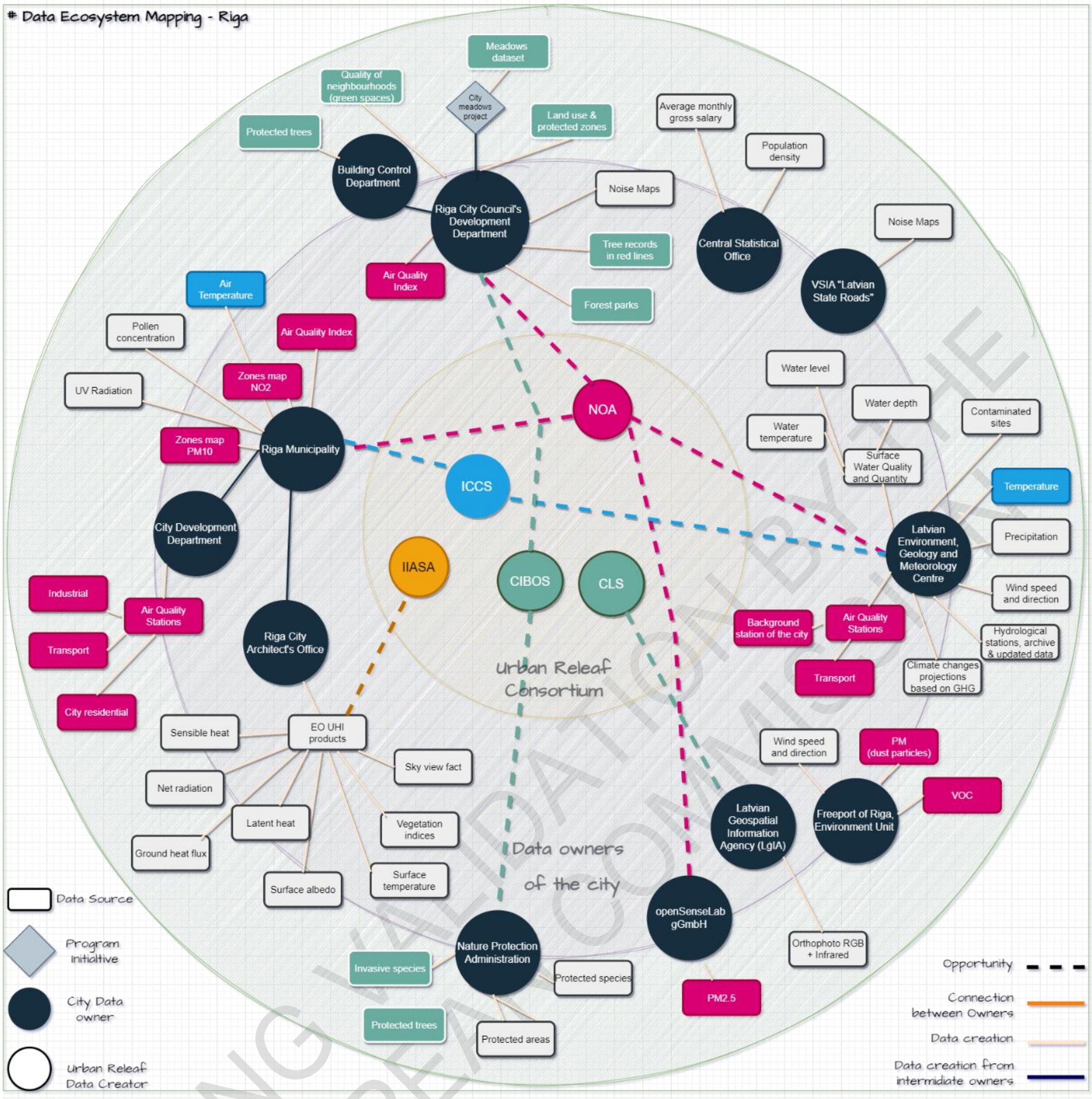





Figure 47: Data Ecosystem Map for Riga

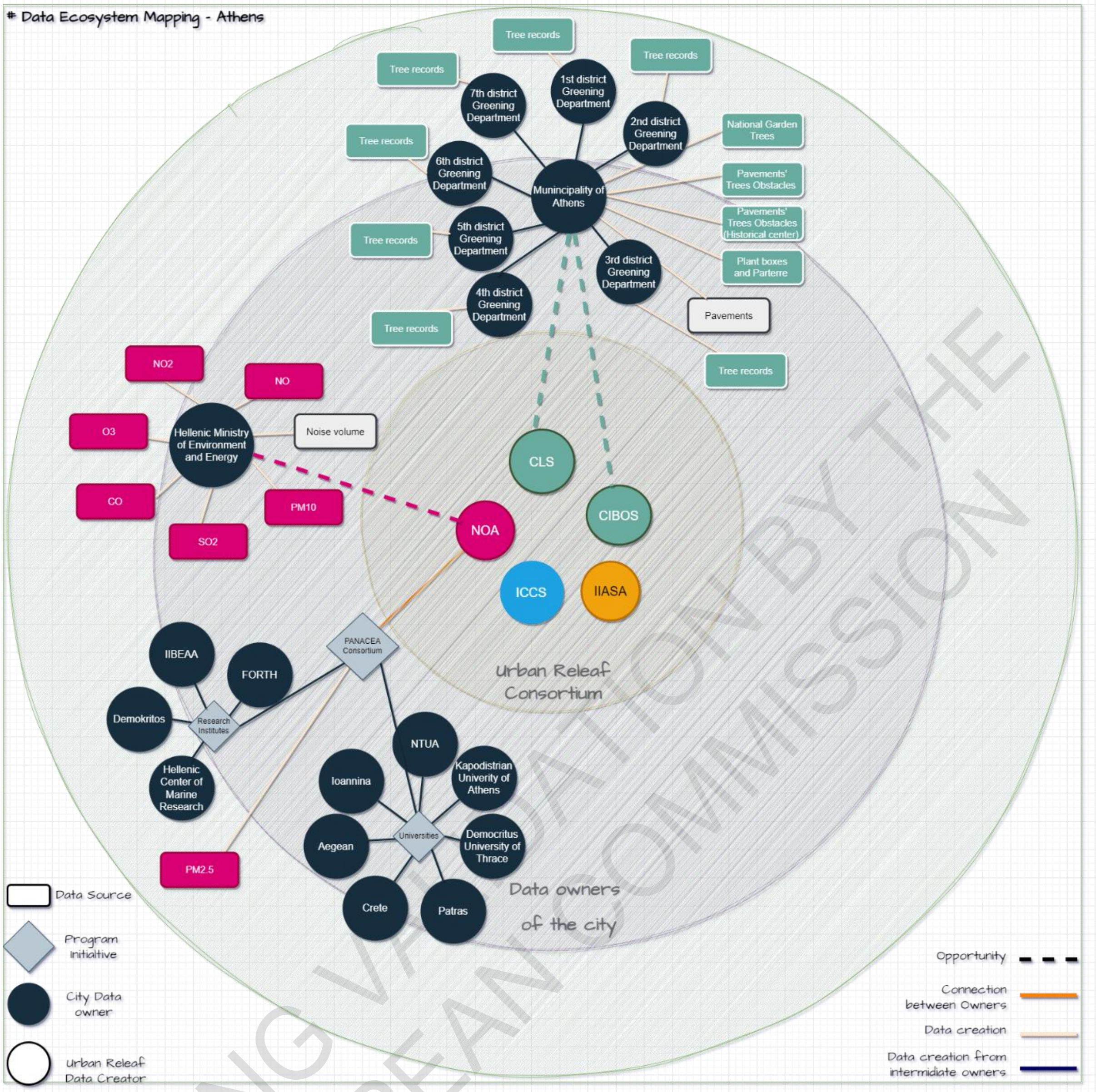



Figure 48: Data Ecosystem Map for Athens

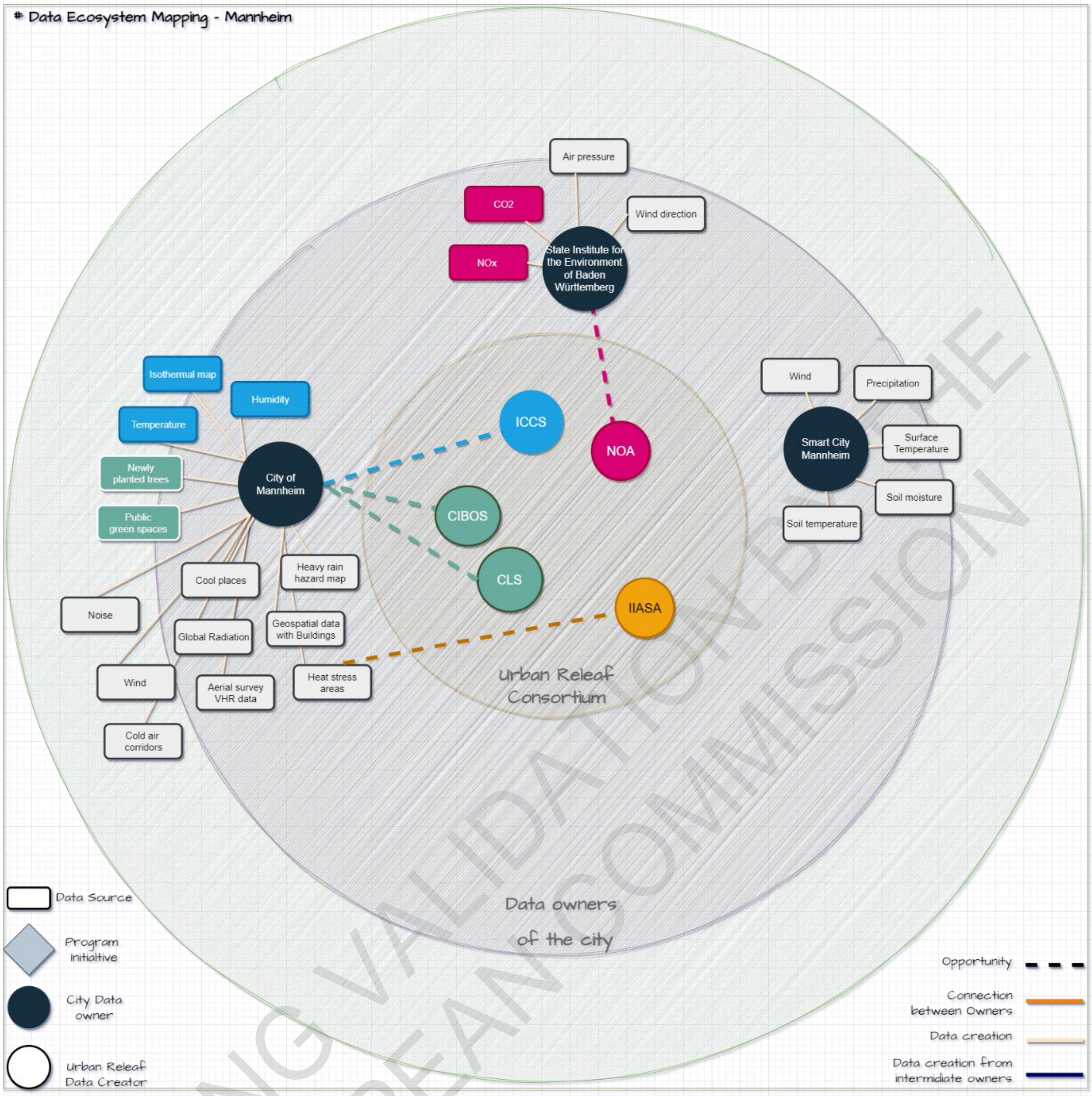



Figure 49: Data Ecosystem Mapping for Mannheim

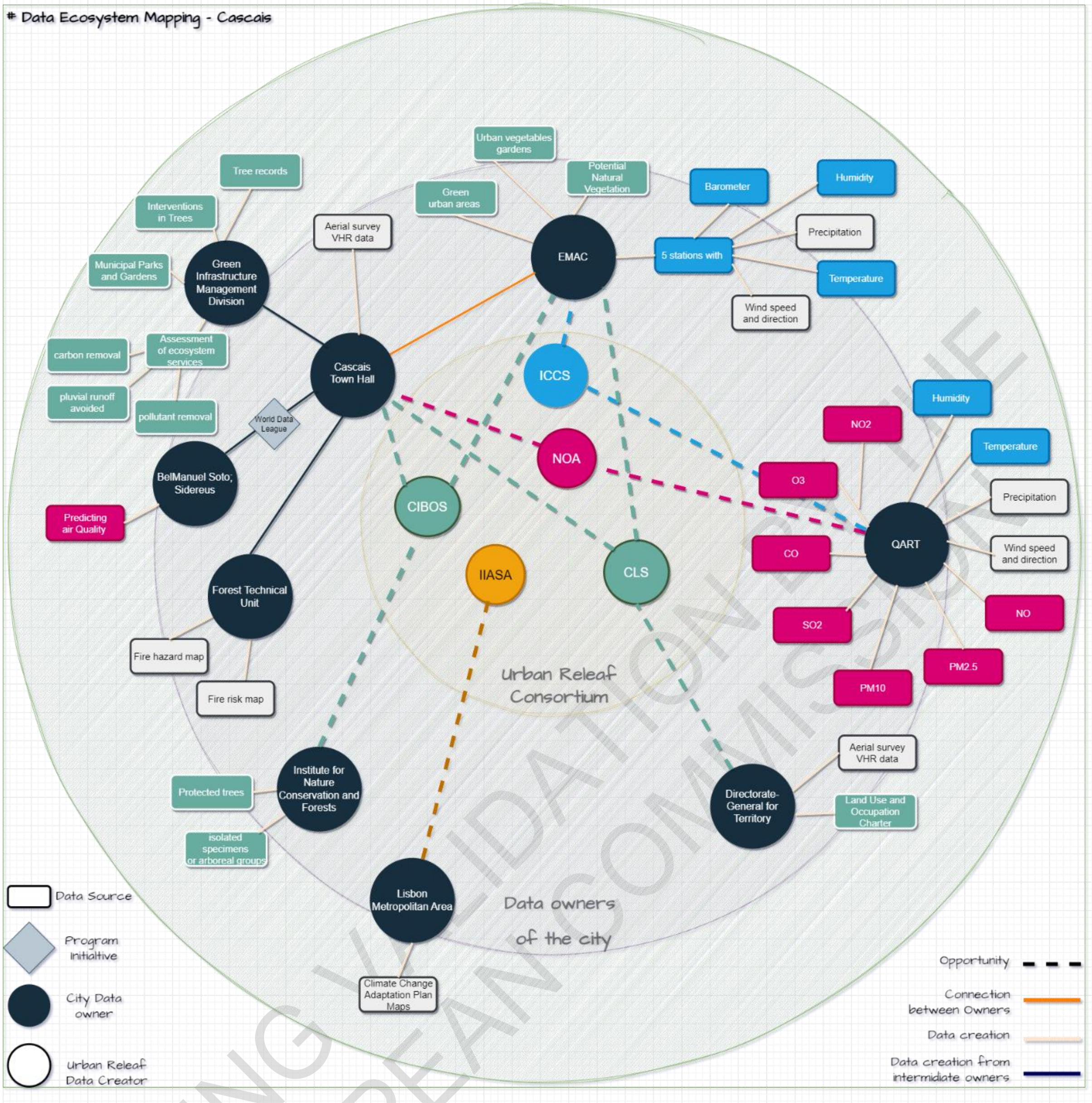



Figure 50: Data Ecosystem Map for Cascais

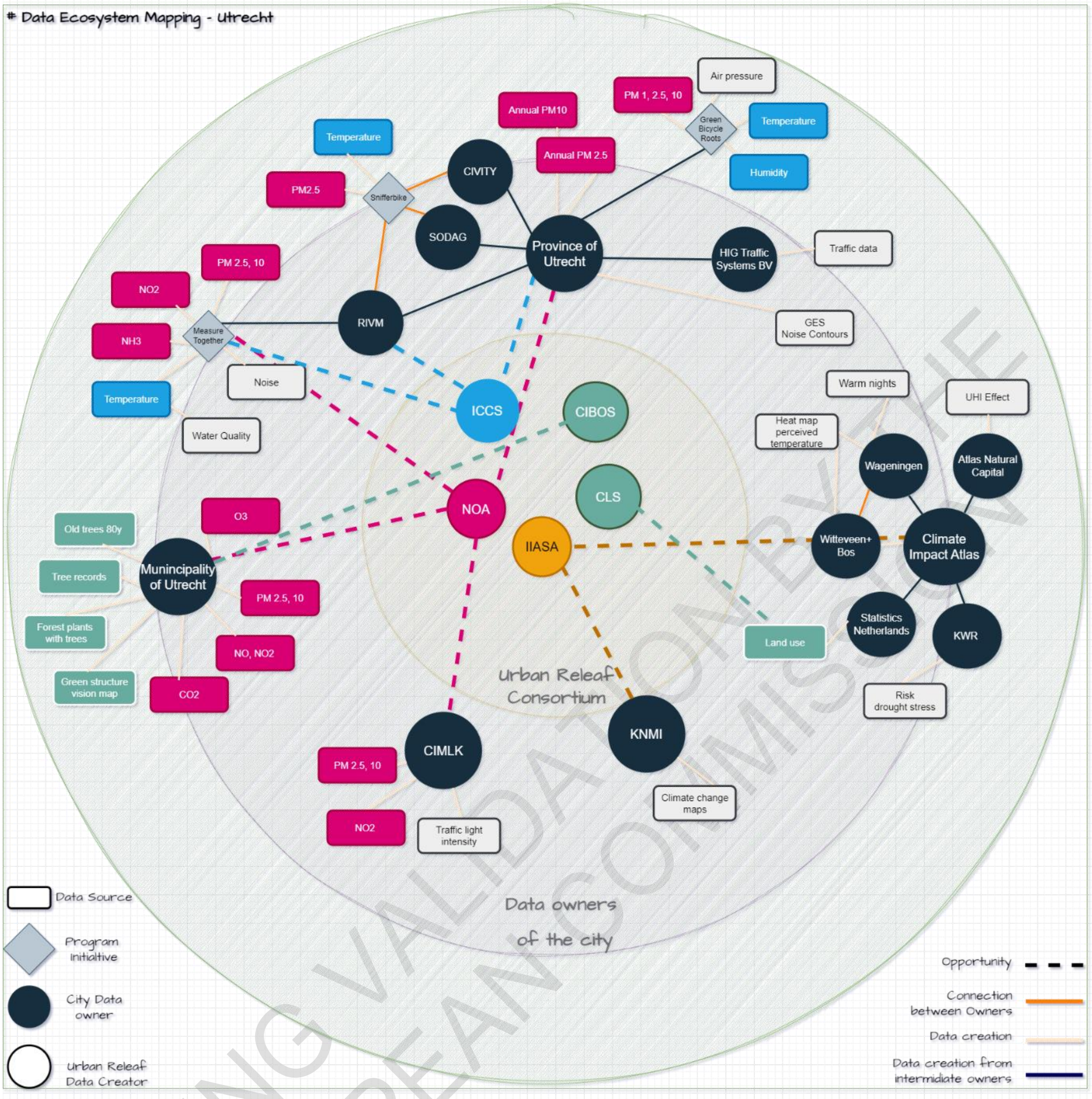





Figure 51: Data Ecosystem Map for Utrecht

## Appendix B: Data Ecosystem Summary

***Data owners** are categorised as the following: **RTI/ACA:** Research Institute & Academic body; **ASSOC/NGO:** Association & Non-Governmental Organisation; **PUB:** Public Administration, **IDN/SME:** Industrial Corporation & Small and medium-sized enterprise.*

| City | Objectives | Pilot City-Specific Goals (i.e. User Stories) | Data owners | Data Assets | |
|---|---|---|---|---|---|
| DUNDEE | **(a)** Improve the existing data sources<br>**(b)** Cope with the data sparsity<br>**(c)** Introduce a “living network of moving observers”<br>**(d)** Open Data Platform (ODP) | **Collect -Tree records and Properties**<br>- Canopy cover, $CO_2$ sequestration, Flood alleviation, Health condition, Categorised as Cultural Heritage Site, Fruit/Nut tree<br>- Mapping trees in private areas<br><br>**Monitor – Areas**<br>TPO areas, Biodiversity monitoring | **PUB**<br>• Dundee City Council (Forestry Department)<br>• Scottish Government (Scottish Natural Heritage, Scottish Forestry)<br>**RTI/ACA**<br>• SEPA<br>• DEFRA<br>• James Hutton Institute<br>• Forest Research<br>**IDN/SME**<br>• Scottish Water | 8 | TPO Boundaries, National Forest (Woodland) Inventory, Tree locations, Native Woodlands and Plantations on Ancient Woodland Sites, Ancient Woodland Inventory, Habitat and Land Cover Map, Local Nature Reserves |
| | | **Collect – Temperature/Humidity CS data**<br>Combine with existing data to produce Heat maps | | 2 | Ambient and Modelled Temperature |
| | | **Collect – PM2.5 CS data**<br>Air Quality Monitoring<br>Well-being monitoring with the combination of passive sampling systems and strava data | | 7 | AQ (single records, maps or projections)<br>• $NO_2$, PM1, PM2.5, PM10, $SO_2$ |
| | | - | | 0 | - |
| | | **Monitor – Areas**<br>Vulnerability assessment<br><br>**Integrate**<br>Connecting the CS data streams into existing platforms of Dundee | | 10 | Wind speed & direction, Noise maps, Open spaces, Soil Inventory, Topsoil Organic Carbon Content (TOC), SPA, SSSI, NVC, RAMSAR, SIMD |
| RIGA | **(a)** T/RH and Air Quality CS sensors | **Collect -Tree records and Properties**<br>Health condition<br><br>**Monitor – Areas**<br>Biodiversity monitoring | **PUB**<br>• Riga City Council development | 9 | Protected areas, Protected species/trees, Contaminated sites, Invasive tree species, EO-spectral indices, Green spaces, Land use and protection zones, Orthophotos, Forests, Tree records, Cities meadows. |

| | | | | | |
|---|---|---|---|---|---|
| | **(b)** Decision-making Data Platform (DDP)<br>**(c)** Monitor the quality of green with CS/EO data<br>**(d)** Citizens' perception and comfort level | **Collect – Temperature/Humidity CS data**<br>Could be an essential addition to the improvement of heat severity maps | department (Building Control department)<br>• Central Statistical office<br>• Riga municipality (City Development department, Riga City Architect's Office)<br>• Latvian Geospatial Information Agency<br>• Nature protection administration<br>**RTI/ACA**<br>• Latvian Environment, Geology and Meteorology Center<br>**IDN/SME**<br>• Freeport of Riga<br>• VSIA "Latvian State Roads" | **2** | Air temperature forecasts, Meteorological operational data (updated and archived) |
| | | **Collect – PM2.5 CS data**<br>Air Quality monitoring-validation-calibration<br>Green space Air quality | | **12** | AQI, AQ PM (dust), VOC (Benzene, NO2, and SO2), SensorCommunity PM 2.5, AQ 3 stations, Monthly AQ reports |
| | | - | | **0** | - |
| | | **Collect – Citizens' perceptions in the vicinity of green territories**<br><br>**Integrate**<br>Connecting the CS data streams into existing platforms of Riga | | **27** | Wind direction & speed, Noise pollution, UHI maps, CC scenarios, Total precipitation, UV radiation, Pollen Concentration, Contaminated sites, Surface water quality, Water level, Water temperature, depth, Population density |
| **ATHENS** | **(a)** Tree inventory<br>**(b)** CS AQ and Temperature-Relative Humidity data<br>**(c)** up-to-date VHR urban greening layer | **Collect - Tree registry**<br>- Comprehensive tree inventory<br>- Tree app/platform<br>- Map trees in private spaces<br>- VHR green layer<br><br>**Monitor – Areas**<br>- Connecting the green spaces<br>- Provide updated greening content facilitating Green management | **PUB**<br>• Greening departments of 7 districts<br>• Municipality of Athens<br>• YPEN<br>**RTI/ACA**<br>• PANACEA consortium<br>**IDN/SME** | **15** | Tree registry, Trees records National Garden of Athens and maps, Tree records reported as pavement objects (Athens district and historical center), Records with plant boxes and parterre; |
| | | **Collect – Temperature/ Humidity CS data**<br>- Provide a dense community network | | **0** | - |

| | | | | | |
|---|---|---|---|---|---|
| | **(d)** contribution to Copernicus Urban Atlas and Panacea | **Collect – AQ data**<br>- High-Resolution BC and UFP exposure<br>- PM2.5 data collection<br>- Standardised Cal/Val methods<br><br>**Monitor – Areas**<br>- Showcase relations between urban green spaces and the improvement in air quality | | **7** | National Air Pollution Monitoring Network (e.g. PM10, CO, NO, NO2, O3, SO2) PANACEA PM2.5 Sensor Infrastructure (Athens Network) |
| | | - | | **0** | - |
| | | - | | **2** | Noise volume (Lden index), Pavements |
| MANNHEIM | **(a)** Tree inventory in diverse locations<br>**(b)** Homogenised dashboard with all CS data and existing data sources | **Collect - Tree registry**<br>- Comprehensive tree inventory<br>- Tree app/platform<br>- Map private trees<br>- Health condition<br><br>**Monitor – Areas**<br>- Connecting the green spaces | **PUB**<br>• City of Mannheim<br>• Smart City Mannheim<br>**RTI/ACA**<br>• State Institute for the Environment of Baden Württemberg<br>**IDN/SME** | **3** | Aerial 3D laser scanning survey, Public green spaces, Newly planted trees |
| | | **Collect – Temperature/ Humidity CS data**<br>- Standardised Cal/Val methods<br>- Generating a unified network | | **4** | Surface infrared temperature; Isothermal map; temperature, humidity (2 sources) |
| | | - | | **2** | AQ [NOx, CO2] |
| | | - | | **0** | - |
| | | **Integrate**<br>- Connecting the CS data streams into existing platforms of Mannheim | | **9** | Wind speed and direction, Cold air corridors, Heat areas, Buildings, Noise, Soil moisture and temperature, Heavy rain hazard maps, Precipitation, Cool places maps, Air pressure, Global radiation |
| CASCAIS | **(a)** Leverage on existing infrastructure<br>**(b)** CS Temperature-Relative Humidity data<br>**(c)** Bioclimatic comfort and | - | **PUB**<br>• Cascais Town Hall (Green Infrastructure Management Division)<br>• Cascais Town Hall (Forest Technical Unit)<br>• EMAC<br>• Directorate-General for Territory | **14** | Green urban areas, Orthophotos, Land Use, Tree registry (protected, interventions) Potential Natural Vegetation, Assessment of ecosystem services [carbon removal, pluvial runoff avoided, pollutant removal] |
| | | **Collect – Temperature/ Humidity CS data**<br>- Provide a dense "alive" community network | | **5** | 16 stations [Humidity, Temperature] |
| | | **-** | | **8** | 11 AQ stations [NO2, O3, PM2.5, PM10, SO2, CO], AQ for Outdoor Activities |

| | | | | | |
|---|---|---|---|---|---|
| | perception measurements over UHI effect | **Collect – Thermal comfort data**<br>- Measuring comfort level over UHI | **RTI/ACA**<br>• Institute for Nature Conservation and Forests<br>**IDN/SME**<br>• QART<br>• Individual scientists | **3** | Climate Change Adaptation Plan, Fire hazard map, Fire risk map |
| | | **Integrate**<br>- Connecting the CS data streams into existing platforms of Cascais | | **4** | 11 AQ stations [noise, pressure, wind speed and direction, precipitation], 5 stations [Barometer, Evapotranspiration, Precipitation, Wind Speed, Direction, Max, WindSamps, Solar Radiation, UV] |
| **UTRECHT** | **(a)** CS Temperature-Relative Humidity data<br>**(b)** Data collection in urban spaces with different profiles<br>**(c)** Integrate data into Utrecht's data platforms | - | **PUB**<br>• Municipality of Utrecht<br>• Province of Utrecht<br>• Scottish Government<br>• Province of Utrecht (HIG Traffic Systems BV)<br>• Statistics Netherlands<br>• Atlas Natural Capital<br>**RTI/ACA**<br>• RIVM<br>• CIMLK<br>• KNMI<br>• Witteveen+Bos<br>• Wageningen<br>• KWR Water Research Institute<br>**IDN/SME**<br>• SODAQ<br>• Civity<br>• Climate Impact Atlas | **5** | Treemap (species, year, tree number), Old trees 80 years and older, Forest plants with trees, Green structure vision map, Land use |
| | | **Collect – Temperature/ Humidity CS data**<br>- The perceived temperature in private gardens<br>- Perceived temperature observations integrated into heat stress models<br>- Accurate air temperature and relative humidity observations | | **4** | Temperature (3 sources), Humidity (1 source) |
| | | - | | **15** | Annual means and single records,<br>- PM2.5 (5 sources), PM10 (2 sources), PM1 (1 source),<br>- AQ pollutants [NO2, CO2, NO, NO2, O3, NO2, NH3] |
| | | **Collect – Thermal comfort data**<br>- Bioclimatic comfort monitoring<br><br>**Monitor- Area**<br>**-** Effect of Green spaces on UHI (combination of T/RH and thermal comfort data) | | **5** | Climate change maps (Number of tropical days, summery days, warm days), Warm nights, Heat maps (2050 Best and Worse-case CC scenario), Risk drought stress, Drought susceptibility nature, UHI |
| | | **Integrate**<br>- Connecting the CS data streams into existing platforms of Utrecht | | **6** | Green Bicycle roots (air pressure); Traffic light intensity; Traffic data (bike, car, truck, pedestrian); Noise, Water Quality |