Horizon Europe

Project: 101138405 — Net4Cities

Deliverable D3.1:
Existing data available in data hub

Net4Cities Consortium

Imprint

Suggested citation:
Vanherle K., Ødegård, R.A., Setia, H. 2025 – NET4Cities Existing data available in data hub. Horizon Europe Project Net4Cities.

Main findings and deliverables of the Net4Cities project will be available at www.net4cities.eu

This project has received funding from the European Union's Horizon Europe funding programme under the call HORIZON-CL5-2023-D5-01 – No. 101138405

Document Control Information

Settings	Value
Document Title:	Existing data available in data hub
Project Title:	Net4Cities
Document Author:	Kris Vanherle (Telraam), Rune Åvar Ødegård (NILU), Himanshu Setia (RIFS)
Doc. Version:	1.3
Sensitivity:	Public, Basic, High
Date:	08/07/2025

Document Approver(s) and Reviewer(s):

NOTE: All Approvers are required. Records of each approver must be maintained. All Reviewers in the list are considered required unless explicitly listed as Optional.

Name	Role	Action	Date
Roberto Sanz	WP Lead	Review	25.03.2025
Erika von Schneidemesser	Coordinator	Review and Approve	31.03.2025
Erika von Schneidemesser	Coordinator	Review and Approve	08.07.2025

Document history:

The Document Author is authorized to make the following types of changes to the document without requiring that the document be re-approved:

- Editorial, formatting, and spelling
- Clarification

To request a change to this document, contact the Document Author or Owner.
Changes to this document are summarized in the following table in reverse chronological order (latest version first).

Revision	Date	Created by	Short Description of Changes
1.3	08/07/2025	Kris Vanherle	Addition of section impact on project

Configuration Management: Document Location

The latest version of this controlled document is stored in the Sync&Share file-sharing platform at “Net4Cities/WP3/D3.1 existing data available in data hub”.

Table of Contents

Executive Summary

The Net4Cities project is a transformative initiative dedicated to enhancing urban air and noise pollution monitoring to support the EU's Zero Pollution Action Plan. As part of Work Package 3 (WP3) on **Noise and Transport Activity Monitoring and Complementary Data Collection**, Deliverable D3.1 focusses on integrating existing traffic count and noise data into the Net4Cities Data Hub, insofar available via easily accessible API's.

This deliverable is a demonstrator (DEM), but for purpose of logging action takes, this document provides a comprehensive assessment of identified available datasets from partner cities, consolidating **supplementary** existing noise and air pollution monitoring infrastructure, citizen science sensor networks, and transport activity data sources. By centralizing this data within the Net4Cities Data Hub, the project enhances accessibility and interoperability, facilitating further use in the project by the project researchers.

We first provide an overview of data sources identified and investigated, to end with selected sources incorporated in the Datahub, including screenshots to demonstrate the actual data ingestion.

1. Criteria for Supplementary Data

To ensure the reliability, consistency, and usability of supplementary data sources for the Net4Cities project, we adhered to the following criteria when identifying potential supplementary data sources for integrating into the project's data hub:

1. **Data Availability & Accessibility**
   - The data must be available through an API or an easily extractable format (e.g., JSON, CSV).
   - Preference for real-time or near-real-time data streams to support timely decision-making.
2. **Data Quality & Accuracy**
   - Sources must provide well-documented methodologies for data collection and validation.
   - The accuracy of measurements should align with European environmental standards.
3. **Data Relevance & Scope**
   - The data should align with Net4Cities objectives, primarily focusing on noise pollution, air quality, and transport emissions.
   - Geographical coverage should be relevant to participating pilot cities.
4. **Interoperability & Integration Feasibility**
   - Data format should be compatible with the Net4Cities Data Hub.
   - Availability of clear API documentation to facilitate integration.
5. **Legal & Licensing Compliance**
   - The data should be open access or come with clear usage rights.
   - Licensing should allow for research, analysis, and potential dissemination.
6. **Sustainability & Long-Term Usability**
   - The data source should have a stable operational history and commitment to ongoing data provision.
   - Preference for sources with institutional backing (e.g., government, research institutions).

2. Overview of Data Sources Identified and Investigated

We used an internal survey to identify data sources, after which these were investigated, using the criteria above. The main focus was usability (i.e. documented API availability and compatibility with the DataHub) and relevance for the project. We used a 3-point scale on each to simplify the selection.

The table below summarizes the data sources reviewed for integration into the Net4Cities Data Hub:

Tabelle 1: list of identified supplementary data sources

Data Source	Type	API Availability	Relevance	Description
Telraam	Traffic counts	Yes	High	Provides segment-based real-time traffic data through API. Includes traffic volume, vehicle types, and speeds. API Documentation
Sensor.community	Air Quality	Yes	High	Open-source community-driven air quality monitoring network. Offers real-time and historical data on PM2.5, PM10, temperature, and humidity. API Documentation
Zürich - Ostluft	Air Quality	No	Low	Provides air quality measurements but lacks API support. Data is available for download in CSV format. Data Portal
Zürich - NABEL	Air Quality	No	Low	National air quality monitoring network with CSV downloads available, but no API for integration. Data Portal
Barcelona Sensors & Actuators	Noise & Traffic	Yes	Medium	Urban monitoring platform with a well-documented API. Requires additional validation for third-party data access. API Documentation
Barcelona Air Quality	Air Quality	Unclear	Medium	Interface available for filtering air quality data, but full API documentation is not publicly available. Data Portal
EEA Strategic Noise Maps	Noise	Yes	Low	Contains aggregated noise pollution data from multiple sources (airports, industries, railways, roads). Useful for understanding the long-term

				acoustic situation in each city, but not for real time noise maps. Data Portal
Tbilisi National Environmental Agency	Air Quality	No	Medium	National environmental monitoring system with downloadable reports in CSV format. Scraping would be required for integration. Data Portal
Southampton AURN	Air Quality	Limited	Low	Part of the UK's national air quality network. Lacks a dedicated API for easy integration. Data Portal
Zephyr Air Quality (Southampton)	Air Quality	Yes	High	High-resolution air quality sensor network. Already integrated into Net4Cities through EarthSense API. Zephyr Information
Limassol Traffic Sensors	Traffic	Limited	High	Provides real-time traffic monitoring data with API support. Suitable for correlation with noise and air quality data. API Documentation
Zürich - Verkehrszählda ten	Traffic	Yes	Medium	Traffic counts on individual streets as separate dataset. No documentation in English. Data Portal

After an extensive review of potential data sources, we decided to proceed primarily with the ingestion of Telraam data for Deliverable D3.1. This decision is based on several key factors:

1. **Complexity of Data Ingestion** – Many of the surveyed data sources either lack an API or have poorly documented and unreliable endpoints. Several API links were found to be broken or insufficiently maintained, making integration infeasible.
2. **Constraints in the Data Hub Architecture** – The Net4Cities Data Hub has specific requirements for georeferenced data ingestion (lat/long or area-based). This rigidity means that additional conversion steps would be required for most data sources, further complicating integration.
3. **Resource Limitations –** Discussions with technical stakeholders, indicate that there is limited capacity to support additional data conversion and ingestion efforts. Given these constraints, focusing on a single, well-documented source like Telraam ensures that efforts remain manageable within the available time and resources.
4. **Reliability & API Quality** – Telraam offers a well-structured, stable API with real-time data availability, making it the most viable option for seamless integration into the Net4Cities Data Hub.

Given these considerations, this deliverable serves as documentation of work done to the **identified supplementary data sources** rather than actual ingestion of multiple datasets in the datahub, which will be limited to Telraam data only.

If, at a later stage, we determine that the sources would be useful for any of the analyses or other products within N4C, we will take the time to ingest the data into the DataHub.

Additionally, data identified as useful in this deliverable, may still be utilized further in the project by directly accessing the available resources, without the need for ingestion into the DataHub.

3. Telraam data schema


```
{
  status_code          integer
                       example: 200
  message              string
                       example: ok
  report               [ {
    instance_id                  integer
                                 example: -1
                                 The instance identifier for "instance" level calls ("-1" for "segment" level calls).
    segment_id                   integer
                                 example: 9000001463
                                 The segment identifier for "segment" level calls (when "instance" calls are implemented, this will read "-1").
    date                         string($date-time)
                                 example: 2024-06-12T11:00:00.000Z
                                 ISO timeflag (date and UTC time) of the reporting interval (beginning of the interval).
    interval                     string
                                 example: hourly
                                 Will be "hourly", standing for hourly aggregate data.
    uptime                       number
                                 example: 0.9995835649
                                 Between 0 and 1, represents the portion of the reporting interval (hour or day) that was actively spent
                                 counting the traffic (devices being offline, background calculation intervals, and the night time can
                                 contribute to values being less than 1, depending on hardware and firmware versions). For Telraam V1 devices
                                 typically under 0.8, while for Telraam S2 devices usually near 1.
    heavy                        number
                                 example: 51.9986103391
                                 The number of heavy vehicles (historically also called lorries) - including anything larger than a car - during
                                 the given period. Corrected for uptime.
    car                          [...]
    bike                         [...]
    pedestrian                   [...]
    night                        [...]
    heavy_lft                    number
                                 example: 11.4997220678
                                 The number of heavy vehicles in the left direction. Corrected for uptime.
    heavy_rgt                    number
                                 example: 40.4988882713
                                 The number of heavy vehicles in the left direction. Corrected for uptime.
    car_lft                      [...]
    car_rgt                      [...]
    bike_lft                     [...]
    bike_rgt                     [...]
    pedestrian_lft               [...]
    pedestrian_rgt               [...]
    night_lft                    [...]
    night_rgt                    [...]
    direction                    integer
                                 example: 1
                                 An internal consistency value making sure that when multiple cameras on different sides of the street are
                                 aggregated then the left and right directions are handled properly. Always 1 for segment level data.
    car_speed_hist_0to70plus     [
                                 example: List [ 8.6355785838, 12.4352331606, 28.8428324698, 18.4801381693, 11.2262521589, 10.0172711572,
                                 5.872193437, 4.4905008636 ]
                                 The estimated car speed distribution in 10 km/h bins from 0 to 70+ km/h (in percentage of the total 100%).
                                 [...]]
    car_speed_hist_0to120plus    [...]
    timezone                     string
                                 example: Europe/Brussels
                                 The name (TZ identifier) of the time zone where the segment is located, which can be used to convert reported
                                 UTC timestamps to local times.
    v85                          number
                                 example: 51.5
                                 The estimated car speed limit in km/h that 85% of all cars respect (15% of drivers drive faster than this
                                 limit). Just like all other speed related measurements, the accuracy of this value is likely not better than
                                 +/-10%.
  }]
}
```

4. Impact on project

The decision to limit the ingestion of data into the Net4Cities Data Hub does not constitute a blocker for the project's successful execution. The core datasets essential to the project—specifically, air quality, traffic, and noise sensor data—have already been ingested and are fully operational within the Data Hub. These data streams form the backbone of Net4Cities' analytical and monitoring objectives and are therefore sufficient to support key project activities and deliverables.

While additional data sources (e.g. municipal traffic counts, low cost air quality networks) were identified and assessed, their integration was deemed non-essential at this stage due to technical constraints, limited API access, or resource limitations. These supplementary datasets, while "nice to have," are not mission-critical and their absence in the Data Hub will not compromise the project outcomes.

Importantly, data not currently ingested can still be used for analysis later in the project. The Data Hub primarily serves to facilitate access and interoperability, particularly for dashboard and Digital Studio use. However, many relevant datasets—such as traffic data from Zurich's Rosengartenbrücke (link)—are openly accessible and can be directly utilized by the research team outside of the Data Hub infrastructure when needed. If deemed sufficiently interesting and feasible to execute, we may still opt to ingest some datasets in the Data Hub later on.

In conclusion, the limited ingestion strategy is a pragmatic decision that balances technical feasibility with project needs, ensuring essential functionality without jeopardizing analytical capability or deliverable quality.

Funded by
the European Union