



Air quality data set evaluation template

SR601



Introduction

This OPENAIR air quality data set evaluation template is intended to be a practical tool to help you evaluate and characterise the quality and usability of data sets linked to your air quality monitoring initiative. This includes data from smart low-cost air quality sensing networks, as well as data from other relevant sources (such as location data, traffic counts, or meteorology).

This template will help you gain a better understanding of your data, by prompting you to consider a series of questions about your data sets:

- What data are you producing or accessing in your project?
- What are the potential uses and limitations of that data?
- What are the data sharing and security considerations?
- What supporting information about your data is needed?
- What could you do to make your data more usable?

The template also serves a secondary role. Once completed, it creates a descriptive data set 'scorecard', providing prospective users with a summary of the collected data, and answering high-level questions they might have about its potential usefulness.

Who is this resource for?

This template is intended for use by project managers and other people who are directly responsible for operating a local government air quality sensing network. The data summary (about each data set) in this template will also be useful to new users and collaborators who may be interested in accessing your project's data.

How to use this resource

This template will guide you in assessing and summarising the utility of your data. Normally, this would take place after an initial data collection period, but it can be beneficial to consider this issue much earlier in the project development process. This helps you to justify continued data collection (e.g. to support internal applications for funding), and to create the basis for new partnerships.

To complete the template, you will be prompted to answer questions about the data you are collecting, such as:

- Where is it from?
- How is it being collected?
- How frequently is it being collected?
- Who is collecting it?

By completing this template, you can provide a simple and reliable overview of your data assets, and assess how well your data sets meet the needs of your planned data use case.



CASE STUDY – DATA USE SCENARIO

A local government agency has established a small trial network of particulate air quality sensing devices.

The planning and placemaking teams have been engaged throughout the design of the project, and data has been collected to explore specific questions identified by these teams. The hope is that an upcoming review of the local government’s Development Control Plan might incorporate insights from the new data.

The project has now collected six months of data, and the planning team is keen to understand more about the data and what they can learn from it. You are the project officer leading the air quality monitoring trial, and you now need to report back to the planning team and to senior management. The success of the trial, and its usefulness to the planning team, relies largely upon the utility of the data that the trial has generated.

By completing this data set evaluation template, you can create a simple and reliable overview of your data assets, and an assessment of how well they can meet the needs of your planned data use case.

You present the completed template at an action meeting with the planning team and your senior managers, and share it with other project stakeholders.

After positive initial feedback from collaborators, you also start to hear of interest from other departments and teams not originally considered as potential data users. You circulate and promote the completed template through internal local government communications channels, and make it widely available. You are subsequently contacted by the local government parks manager, who is interested in using the data you are collecting to support a new green infrastructure initiative currently being developed. This leads to a new ongoing internal partnerships and new opportunities to provide value to your local community.

The air quality data set evaluation template

Any data set can be characterised using metadata fields or ‘indicators’, which are listed in the first column of this template (see Table 1).

Table 1. Guide to assessing each data set

Indicator	Description of indicator	Details of what to include
Data set name	Title of the data set	Air quality monitoring data
Data description	Characterises the data	<ul style="list-style-type: none"> What sort of data is this? <i>Particulate matter (PM_{2.5})</i> <i>Description:</i> PM_{2.5} is the concentration of suspended particles (measuring ≤2.5 microns in diameter) within a volume of air.
Data collection	Characterises the way in which data is collected	<ul style="list-style-type: none"> How, where, and when has the data been collected? <p>Include the following details for sensing data:</p> <ul style="list-style-type: none"> device make/model number of devices logging/sampling interval communications type (e.g. Wi-Fi, LoRaWAN) short description of deployment methodology time frame – start and end dates locations (consider including photographs of locations). <p>Include details about any other kinds of related data sets (e.g. traffic information, planning information, and weather):</p> <ul style="list-style-type: none"> Who provided the data? How was it collected? How often will it be updated?
Data format	Characterises how data is stored and can be accessed	<ul style="list-style-type: none"> Is your data a live stream or static file? What API or file format is it in (e.g. JSON, CSV, PDF)? Provide a link to a sample of the API data, if available.

Indicator	Description of indicator	Details of what to include
Data validation	Characterises the quality of the data	<p>For sensing data:</p> <ul style="list-style-type: none"> • What is the date of last calibration? • Have any correction factors been applied? If yes, which ones? • Have sensing devices been co-located with reference sensing devices? • Is the data 'raw', or has it been processed in some way (e.g. by applying statistical analysis like averaging)? • Is there is a quality assurance process in place for the data? • Is the data set complete, or is there known to be missing data?
Data access and sharing	Characterises the decisions that have been made about how the data can be shared	<p>Data storage and access:</p> <ul style="list-style-type: none"> • Where is the data stored? • Who will have access to the data (internal and external users)? • Is sharing of the data covered by an organisational data sharing policy? • How will you and others access the data? • If the data is external, what are the licensing conditions and costs? • Is there a data management and storage plan? • Who owns the data? • Who is responsible for keeping the data up to date? • What decisions need to be made around copyright and intellectual property? • What are the identified data access roles and responsibilities?
People	Characterises the roles and responsibilities relating to management of (and access to) the data resource	<p>Primary data owner (within local government):</p> <ul style="list-style-type: none"> • Name • Contact details <p>Secondary point of contact (within local government):</p> <ul style="list-style-type: none"> • Name • Contact details

Further information

For more information about this project, please contact:

Peter Runcie

Project Lead, NSW Smart Sensing Network

Email: peter@natirar.com.au

This supplementary resource is part of a suite of resources designed to support local government action on air quality through the use of smart low-cost sensing technologies. It is the first Australian project of its kind. Visit www.openair.org.au for more information.

OPENAIR is made possible by the NSW Government's Smart Places Acceleration Program.

Document No: 20231127 SR601 Air quality data set evaluation template Version 1 Final

