

Best Practice Guide

BP403 | Manage and analyse data

Sharing air quality data



Introduction

Until recently, affordable air quality sensing technology has generally not included real-time data connectivity. Collecting data from air quality sensors tended to require manual techniques, such as retrieving memory cards, or downloading data from sensing devices on-site (using a wired connection).

With the advent of the Internet of Things (IoT) and smart cities in the last decade, local governments now have access to relatively low-cost technologies that directly connect small sensing devices with data users. Devices can now transmit air quality measurements wirelessly, every few minutes. This has fundamentally altered the way environmental sensing is approached at a local level, and opens the door to a wide range of new data-driven use cases.

Many of these new use cases require sharing the project data within organisations, or with external users. These users may include universities, commercial partners, government agencies, and the general public.

This creates a new set of challenges for data custodians. An air quality monitoring project will generate a lot of data, and it is vital that this data is shared safely, securely, and in a way that is 'fit-for-purpose' for data users.

Who is this resource for?

While this chapter contains some high-level technical information, it is intended for general use by all parties involved in a local government air quality sensing project, including:

- people leading new air quality monitoring projects
- smart city professionals
- planners
- environmental officers
- local government leadership
- information, communication and technology professionals data custodians
- analysts.

How to use this resource

This Best Practice Guide chapter introduces concepts and best practice approaches for sharing data in general (and for sharing air quality monitoring data specifically). It includes a checklist of factors to consider when sharing air quality data.

Figure 1 maps a series of OPENAIR resources on data policy and data sharing, and how they are related. It is recommended you first read and understand *Data policy for local government air quality monitoring*, which establishes data policy as the critical foundation for best practice data sharing (focusing on the 'why'). Next, engage with this current Best Practice Guide chapter (*Sharing air quality data*), which focuses on the 'what'.

The factsheets *Data policy for local government air quality monitoring* and *Sharing air quality data* provide quick summaries of each topic.

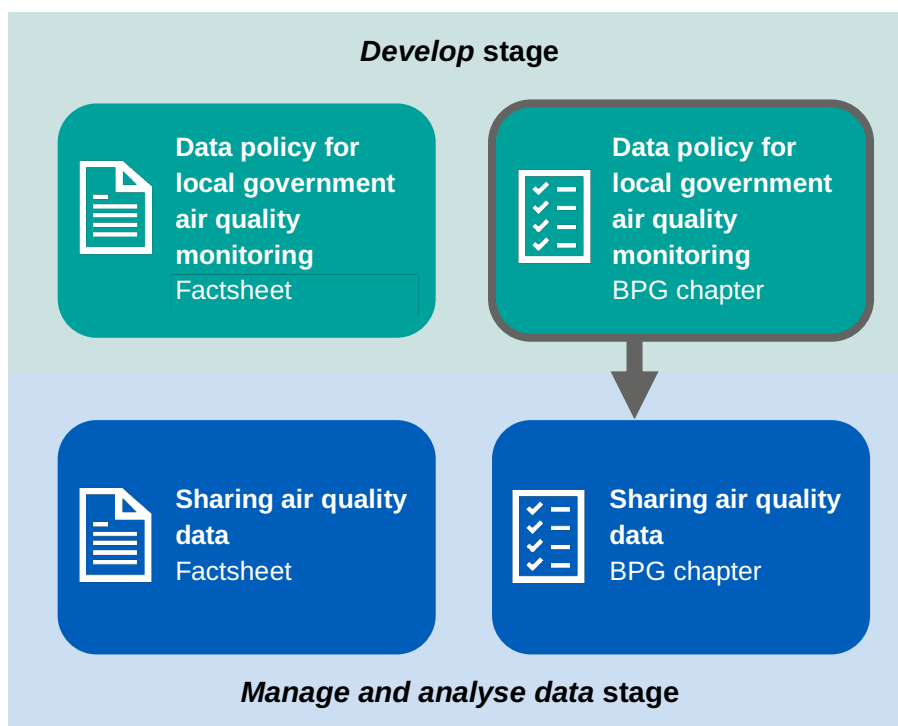


Figure 1. OPENAIR data policy and data sharing resources

Introduction to data sharing

Any local government that is beginning to collect air quality data (or any smart city-related data) should approach data sharing in an informed and considered way.

The rationale for sharing data should be very clear from the outset. The OPENAIR Impact Planning Cycle is designed to help organisations consider their reasons for embarking on an air quality monitoring project, and to plan for impact from the start (see the Best Practice Guide chapter *The Impact Planning Cycle overview* for more details).

Initially, air quality monitoring data is collected for a specific purpose. The resulting data set may be used to inform an internal environmental study, to engage with another government agency, or to be shared with the local community.

In the broader smart city context, data sets created for one purpose may be used in future for other purposes that reach far beyond the initial identified group of data users. *Open data* is one example of data that can be used for a wide range of purposes, and with full visibility for all users. In this case, it is important that data users understand that the data is being shared, its quality, and other characteristics, so they can determine if it is fit for their purposes.

This chapter discusses the practicalities of data sharing in the context of low-cost air quality monitoring, and covers:

1. Key concepts about data sharing
2. A checklist of factors to consider and implement, as part of your data sharing initiative.

Data concepts

Two key concepts relating to data sharing are:

- The data life cycle
- Data sharing risks.

The data life cycle

Like physical assets, data has a life cycle, and there are costs associated with each stage of its life cycle. Local government assets (such as vehicles) follow a procurement process, servicing regime, and disposal process. The data life cycle is similar.

Data is either purchased or gathered from sensing devices, or from other sources. The sensing devices and associated equipment need to be maintained and operated. Data maintenance requires specialist skills and qualifications, as well as operational budget and resources. When data reaches the end of its useful life, decisions must be made regarding its safe and effective disposal. Unlike with tangible assets, disposal options are not always obvious for data.

Data sharing risks

In the context of government data, the data sharing risk of most concern is personal privacy risk. This is the risk that, by sharing data, some private information about individuals may be disclosed.

Although air quality monitoring data generally presents no privacy concerns, it may be possible to combine it with other data sets in a way that increases the risk of identifying individuals.

For example, temperature data gathered from a private home address may provide a fragment of information to an analyst that discloses when the resident is at home at a specific time. This is personal information, and care must be taken to protect this data. Any data that can be related in any way to an individual may enable analytics to identify that individual.

An introduction to techniques to assess and mitigate privacy risks from data sharing are described in The Australian Computer Society's [Data Sharing Frameworks document](#), published in 2017.

Data sharing checklist

Use the following checklist to ensure you are adequately prepared to share your air quality data in a way that is fit-for-purpose, effective, and addresses potential risks.

1. Ensure your organisation is ready
2. Create data policy
3. Acquire the data
4. Determine what data to share
5. Maintain a data set inventory
6. Assess data sets
7. Share your data
8. Provide training and support
9. Periodic reviews

1. Ensure your organisation is ready

Organisational readiness is about ensuring that organisational structures, job roles, skills, and procedures are in place to support data sharing.

Organisational structures

Data is an important local government asset, similar to local government-owned vehicles, buildings, and other resources. As such, it is important that people and departments who acquire, maintain, and share data are accountable for it, and coordinate adequate oversight and escalation paths.

Each organisation will have its own way of organising staff to undertake and oversee data sharing.

Figure 2 illustrates a virtual team structure for managing shared data. It comprises both local government departments, and relevant partner organisations. Note that for smaller pilot projects, there will likely be only a single data owner, who may also be the data champion¹.

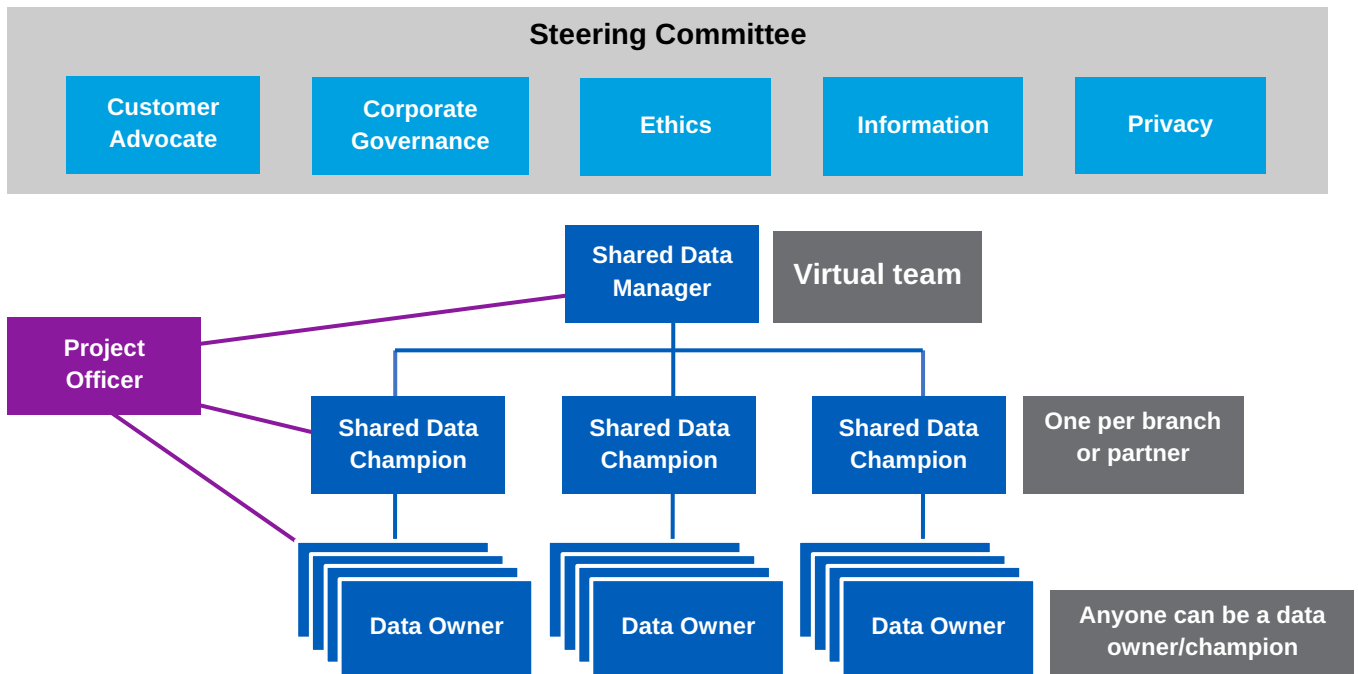


Figure 2. Data sharing accountability – a virtual team structure

The steering committee in this example comprises the manager, along with local government branch managers/partners (from corporate governance, privacy, ethics, and information technology departments). This committee has oversight of the shared data, and is the point of escalation for decision-making support.

Roles and responsibilities

Before sharing data on a routine basis, it is important that your departments and staff are clear about their roles and responsibilities.

Table 1 describes typical job roles and responsibilities that may apply to each role, with respect to data sharing. Use this table as a guide to define these roles within your own organisation.

¹ Data champions are people who advise the way data is handled. (University of Cambridge, n.d.)

Table 1. Data sharing roles and responsibilities

Job role	Data sharing responsibilities
Department manager	<ul style="list-style-type: none"> • Ownership of the business responsibilities for their department, and business outcomes for data sharing by their department. • Member of the data sharing steering committee.
Corporate governance manager	<ul style="list-style-type: none"> • Manage legal and risk matters arising from sharing data, including data privacy risks and ethical considerations. • Member of the data sharing steering committee.
Chief Information Officer (CIO)	<ul style="list-style-type: none"> • Manage technical infrastructure and interoperability. • Member of the data sharing steering committee.
Customer advocacy manager	<ul style="list-style-type: none"> • Represent the voice of the customer (particularly relevant for open data made available to the general public).
Shared data manager	<ul style="list-style-type: none"> • Ensure all data shared by the organisation adheres to the shared data policy, and follows assessment processes and templates. • Hold regular meetings with the shared data champions to provide program updates and guidance, and solicit suggestions for data sharing improvements. • Maintain an up-to-date list of stakeholders, and regularly engage with and include them in project activities. • Track and respond to stakeholder questions and data set requests. • Lead the data set evaluation process, and work with the steering committee on any issues. Develop the final recommendation deliverables and data sharing assessment.
Project manager	<ul style="list-style-type: none"> • Guide and ensure data owners make data (and accompanying metadata) machine-readable and available digitally. • Ensure shared data adheres to local government’s licencing, privacy, security, retention, and public disclosure policies and standards. • Identify a shared data champion in the relevant department, who will serve as the point of contact and coordinator for publishing shared data. • Report to the shared data manager on performance and attainment of data policy goals.

Job role	Data sharing responsibilities
	<ul style="list-style-type: none"> • Ensure any sensitive data is handled by privacy-approved personnel, in a suitable privacy-approved workplace, to avoid unintended and inappropriate sharing. Workplace includes the physical environment, and ICT infrastructure.
<p>Shared data champion</p>	<p>Each department's champion is responsible for managing their department's data sharing. They will work closely with the shared data manager.</p> <ul style="list-style-type: none"> • Gather and maintain a comprehensive inventory of data sets collected, created, or maintained by the department or project, including those managed through contracts with a third party. • Document and publish data creation processes for each data set. • Evaluate the suitability of data sets for publishing. • Prioritise data to be published. • Maintain an inventory of shared data published by the department or project. • Ensure data is produced in a machine-readable form. • Ensure publication of shared data according to defined timelines. • Ensure shared data is accurate and current. • Coordinate the publication of the data with the data team.
<p>Privacy officer</p>	<p>Privacy officers provide the subject matter expertise for public information, and are responsible for understanding privacy issues, and raising them with the project manager to discuss with the steering committee.</p> <ul style="list-style-type: none"> • Serve as arbiter of any questions or issues concerning shared data privacy, or legal solutions for minimising risks of privacy-related harm. • Resolve questions or issues regarding shared data privacy. • Participate in the annual assessment for shared data.

2. Create data policy

The purpose and scope of data policy is described in the OPENAIR Best Practice Guide chapter *Data policy for local government air quality monitoring*.

Policy alignment

Your data policy should be aligned with local government data, security, data retention, public disclosure policies, standards, and (where possible) with state government guidelines for the consistent identification and handling of all shared data.

Scope of data policy

Data sharing policy should ideally be an organisation-wide policy, and not specific to any given project. If your organisation is just starting out on its 'smart city' journey, you can consider restricting the scope of your data policy initially (to one or a few pilot projects).

Privacy and commercial sensitivities

Local government data sharing policy, in general, is to share all data openly, providing it does not create a personal re-identification risk, or have any other form of security risk, commercial restriction, or concern. For these reasons, in some cases, it may necessary to restrict data sharing to a limited number of trusted partners or organisations. This document will help organisations make informed decisions about the sharing of data from air quality monitoring initiatives.

3. Create and maintain a data set inventory

The data set inventory serves as the 'asset register' for your data sets. It will include information about data sources, who has access to the data, the quality of the data, and how the data is maintained.

If your organisation does not have a data set inventory, then you need to create one for your air quality monitoring project. Over time, the inventory should be expanded to include all data sets that local government is sharing, or may share in future.

4. Acquire the data

Deploy air quality monitoring equipment, and start collecting data from your smart low-cost sensing devices.

Arrange for access to licenced and purchased data from third parties, if needed.

5. Determine what data to share

Air quality data (and other environmental monitoring data) consists not just of the air quality measurements, but also contains supporting types of data needed to interpret and maintain the data. Not all of this data needs to be shared, so you need to consider which kinds of data to share. See Table 2 for the four categories of air quality data that will usually be collected in a sensing project.

Table 2. Four categories of data typically collected and stored as part of an air quality monitoring initiative

Category	Description
Environmental telemetry data	<p>Telemetry data consists of the environmental measurements generated by environmental sensing devices. Examples include PM_{2.5}², temperature, humidity, ozone, and NO₂.</p> <p>Telemetry data is updated frequently (often every 10 to 15 minutes, or even more frequently).</p> <p>This is the data that directly supports your business objectives and data use case, and is of most interest.</p>
Operational telemetry data	<p>Many sensing devices will also generate changing data that relates to the operations of the device itself.</p> <p>Examples include power or battery status, wireless communications signal strength, temperature within the sensing enclosure, and health/error information.</p> <p>Typically, this data would be used by IT operations support, and not shared as part of an environmental monitoring program.</p>
Static data	<p>Static data describes aspects of the sensing device, its deployment, and operation. It tends to be created, and then updated only infrequently (if at all).</p> <p>Some of this data may be shared, as it provides useful context for the environmental telemetry data, including location (latitude, longitude, altitude, and/or address) and organisational details.</p>
Metadata	<p>Metadata is data <i>about</i> other data. It provides context to help interpret the data of interest. It also makes finding and working with data easier, facilitating searching and filtering of data sets.</p> <p>Telemetry data sets will also include metadata to help interpret the measurements. Typically, this is the date and time for each measurement, and some form of location information (such as a site identifier).</p> <p>Some of the static data listed above can be considered as metadata, as it helps users understand more about the environmental telemetry data, and how it was collected.</p> <p>It is useful to share some metadata, including information about the sensor sampling rate, resolution, and accuracy.</p>

² PM (particulate matter) refers to airborne solids or liquids. Its size is measured in micrometres and is indicated by the subscript. E.g. PM_{2.5} has a diameter of 2.5 micrometres or less. (NSW Health, 2020)

6. Assess data for sharing

Care must be taken to actively decide when a data set is suitable for sharing, and with whom it can be shared. Without a formalised process, ad hoc decisions can lead to inconsistency, and unforeseen consequences.

Sensitive data – The Five Safes model

A number of organisations around the world (including the Australian Bureau of Statistics) use a model referred to as the ‘Five Safes’. Originally developed by the UK Office of National Statistics, the Five Safes shown in Figure 3 is a framework for helping to make decisions about the effective use of data that may be confidential or sensitive. It was originally used to describe or design research access to statistical data held by government agencies, and by the UK Data Service.

Environmental monitoring data is generally low-risk, and has minimal privacy implications. However, there are circumstances where personal privacy can be compromised.

The Five Safes is a system framework. It is intended to review how all the data sensitivity elements fit together. When embarking on an air quality monitoring project or other initiatives that will result in the creation of new data sets, it is useful to work through the Five Safes method. Ensure care is taken to appropriately handle the data that might be an input to the initiative, as well as the data outputs created. Note that this model deals only with the sensitivity of the data, and any concerns regarding the protection of people’s identity or other privacy matters.

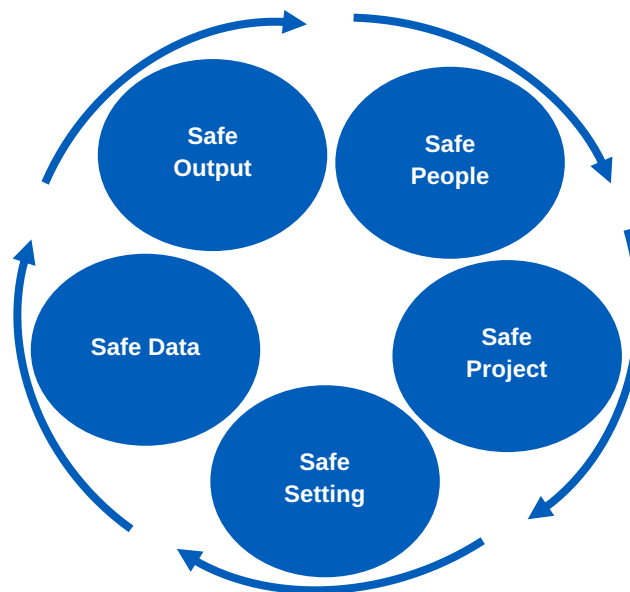


Figure 3. The Five Safes Model

Safe Setting

Refers to the practical controls on the way the data is accessed. At one extreme, researchers may be restricted to using the data in a supervised physical location. At the other extreme, there are no restrictions on data downloaded from the internet.

Safe Projects

Refers to the legal, moral, and ethical considerations surrounding use of the data. This is often specified in regulations or legislation, typically allowing but limiting data use to some form of ‘valid statistical purpose’, and with appropriate ‘public benefit’.

Safe People

Refers to the knowledge, skills, and incentives of the users to store and use the data appropriately. In this context, ‘appropriately’ means ‘in accordance with the required standards of behaviour’, rather than level of statistical skill.

Safe Data

Refers primarily to the potential for identification in the data. It could also refer to the sensitivity of the data itself.

Safe Outputs

Refers to the residual re-identification concerns remaining in sensitive data.

Within the *Safe Data* and *Safe Outputs* elements, there are several specific factors that should be considered, relating to data sharing:

1. **Data Security and Completeness** – checking that all the data is complete, and there is nothing that could be considered a security breach by local government.
2. **Data Complexity and Primacy** – checking that the data actually contains the details needed by the data users, and that it is simple enough to understand and use effectively.
3. **Data Maintenance** – checks that the data set is able to be well-maintained by its owner, and that the data owner has committed to maintain the data to an agreed standard of detail, accuracy, and timeliness.
4. **Data Usability and Licence** – checks that the person receiving the data can read it, and that all the ownership rights and user rights are in place.
5. **Metadata Completeness** – checks that the data can be machine-readable, and that the metadata is complete and contextually appropriate for the user.

7. Share the data

There are many platforms available for sharing data. When selecting and configuring your data sharing platform, consider the following:

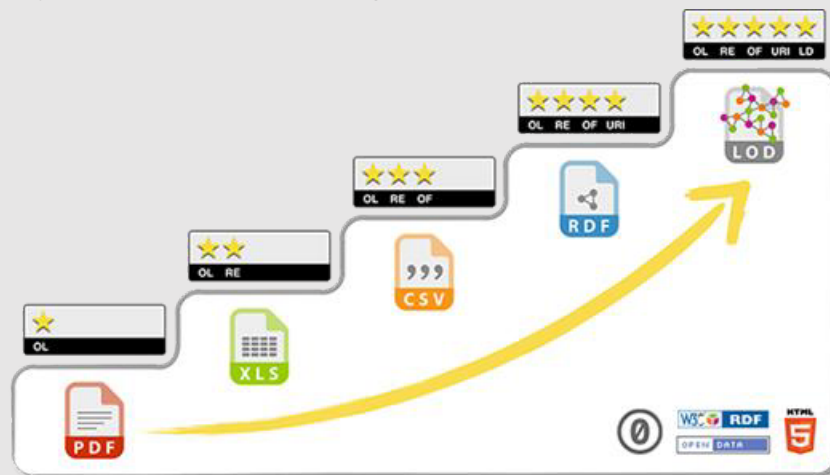
- Configuring access permissions, so that only the intended users can access the data.
- Logging and reporting access – this is useful to understand who is using the data, and how often.
- Data sharing formats – how data is shared will considerably affect its usefulness. Data should ideally be shared in common, easily accessible, machine-readable formats.



5-STAR OPEN DATA

Tim Berners-Lee, inventor of the Web and Linked Data initiator, suggests a 5-star deployment scheme for sharing open data. This applies to all data sharing from trusted partners to the public, for example:

- ★ make data available on the web (any format)
- ★★ make data available as structured data (e.g. Excel instead of image scan of a table)
- ★★★ non-proprietary format (e.g. CSV instead of Excel)
- ★★★★ use URLs to identify data, so users can link to the data
- ★★★★★ link your data to other data to provide context



Best practice is for all shared data to have at least 4 stars, and ideally 5.

8. Provide training and support

Provide appropriate awareness training and support for data champions, and other relevant resources within the data sharing community.

Information about the data should also be provided to data users, including the general public. This should cover how the data was collected, and how to interpret and understand it.

9. Periodic reviews

Plan to review and update your data policy when new projects involving data are proposed, and periodically to ensure the policy is aligned with other local and state government data policies.

Regularly assess the relevance and usefulness of data, as well as its maintenance needs. When the data has no further use and is not being maintained, it should be removed. Ensure any licencing fees are terminated if data will no longer be used or maintained.

References

NSW Health. (2020). *Particulate matter (PM10 and PM2.5)*. NSW Government.

<https://www.health.nsw.gov.au/environment/air/Pages/particulate-matter.aspx>

University of Cambridge. (n.d.). *What is a Data Champion?* [https://www.data.cam.ac.uk/intro-data-champions#:~:text=What%20is%20a%20Data%20Champion,usable%20\(FAIR\)%20research%20principles.](https://www.data.cam.ac.uk/intro-data-champions#:~:text=What%20is%20a%20Data%20Champion,usable%20(FAIR)%20research%20principles.)

[https://www.data.cam.ac.uk/intro-data-champions#:~:text=What%20is%20a%20Data%20Champion,usable%20\(FAIR\)%20research%20principles.](https://www.data.cam.ac.uk/intro-data-champions#:~:text=What%20is%20a%20Data%20Champion,usable%20(FAIR)%20research%20principles.)

Additional resources

Related data sharing policies and legislation

There are many policies and legislation that are relevant to data sharing. Some examples include:

- **Australian Computer Society | [Data Sharing Frameworks](#)**
This is an industry resource.
- **Office of the National Data Commissioner | [Share Data](#)**
This is a government resource.
- **Information and Privacy Commission | [Guide - Data Sharing and Privacy](#)**
This is a government resource.
- **Data.NSW | [Data Sharing Principles](#)**
This is a government resource.

Associated OPENAIR resources

Factsheets

Data policy for local government air quality monitoring

This factsheet provides an overview of how local government data policy can support the effective, responsible and strategic management and sharing of data associated with air quality monitoring.

Sharing air quality data

This factsheet provides an overview of how to share air quality data, including how to decide what data to share, and how to ensure that data is shared appropriately and effectively.

Best Practice Guide chapters

Data policy for local government air quality monitoring

This chapter explores how local government data policy can support the effective, responsible and strategic management, and sharing of data associated with air quality monitoring. It addresses critical considerations and provides practical advice relating to the design and development of appropriate data policy.

Sharing air quality data

This Best Practice Guide chapter provides guidance on the sharing of air quality data. It explores the process by which a local government might assess data to determine its shareability, and presents a series of practical options for implementing data sharing.

Engaging your community with air quality data

This Best Practice Guide chapter provides guidance for engagement of community with air quality data. It explores the benefits of community data activation, provides tips on how to get started, and presents an extensive range of community engagement approaches that a local government might choose to adopt. It ends with a look at how to mitigate common risks associated with data release and community engagement with data.

The Impact Planning Cycle overview

This Best Practice Guide chapter introduces the OPENAIR Impact Planning Cycle, a simple, practical framework designed to assist local governments with impact planning for a smart air quality monitoring project. The Impact Planning Cycle is a planning tool that can help to maximise the impact of a project, and address the needs of an organisation and community.

Further information

For more information about this project, please contact:

Peter Runcie

Project Lead, NSW Smart Sensing Network (NSSN)

Email: peter@natirar.com.au

This Best Practice Guide chapter 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: 20231207 BP403 Sharing air quality data Version 1 Final

