OPENAIR

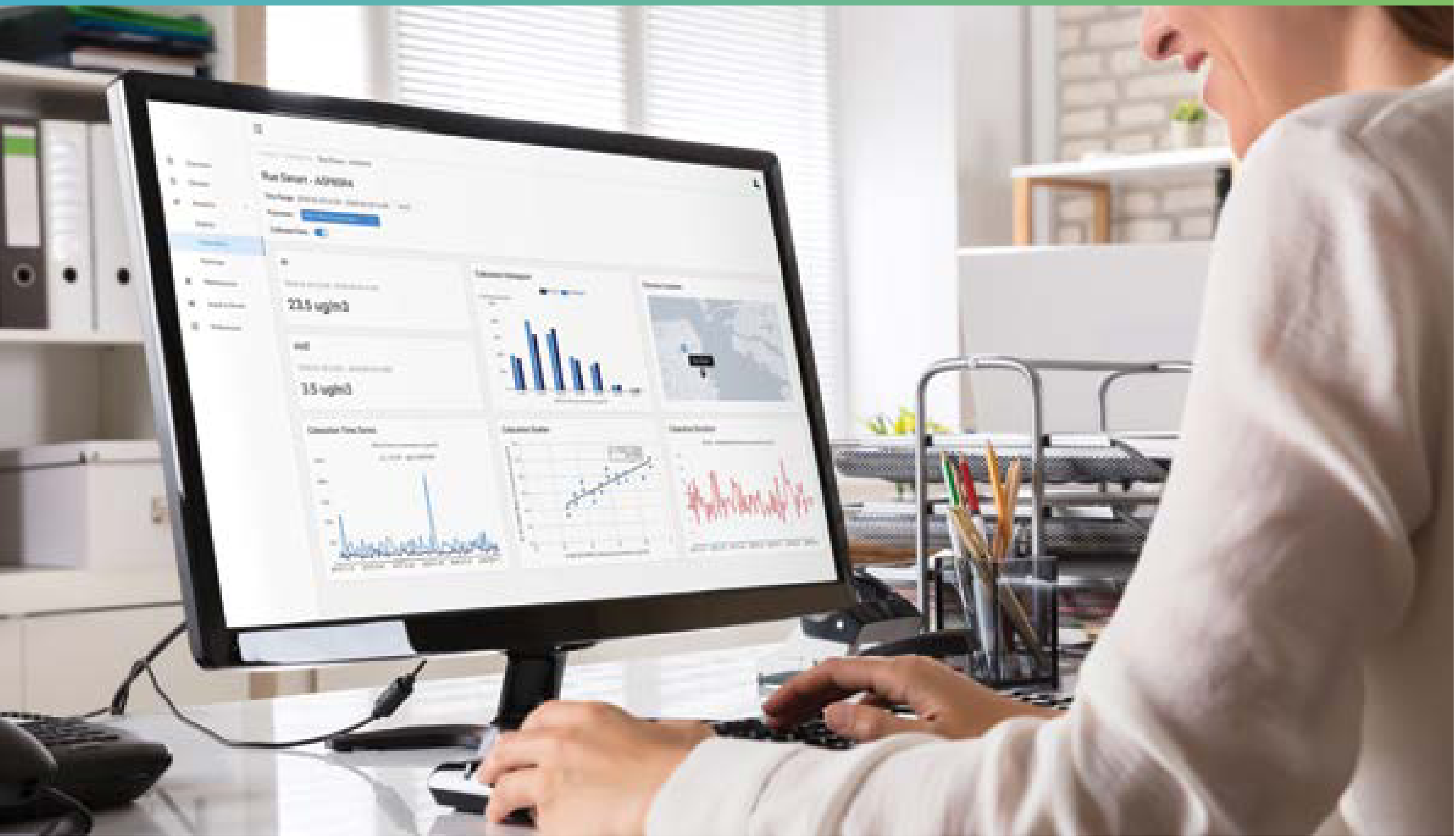


Image source: courtesy of Clarity Movement Co

Air quality data interpretation

Air quality data from smart low-cost air quality sensing devices has the potential to support a wide variety of high-impact activities.

However, raw data on its own has limited use. To extract value from raw data, it must be interpreted to make it useable and able to produce insights or functional outputs that users can trust as being accurate.

The potential value of data therefore depends on the capacity of an organisation to undertake effective data interpretation, and to act on the insights that emerge.

How does data interpretation support impact?

The effective interpretation of raw data can produce insights about a particular issue, or functional outputs (such as automated alerts), both of which can then drive activities for impact creation.

Raw Sensor data → **STEP 1** Data correction & harmonisation → **STEP 2** Data quality control → **STEP 3** Data analysis → Insights or Functional Outputs → Activities for impact

Top tips for supporting effective data interpretation:

1. **Have clear project objectives** that connect a problem statement with fit-for-purpose data, data users, and an impact plan.
2. **Make sure that your sensing devices can produce data with the attributes required** by your data interpretation approach (e.g. reliable and complete data sets, or high levels of accuracy).
3. **Clearly define a technical solution** that is capable of supporting data interpretation that meets your impact objectives.
4. **Assign responsibility** for data interpretation tasks.
5. **Allocate time and budget** to data interpretation tasks as part of your project plan.

Manual vs automated approaches to data interpretation

MANUAL DATA INTERPRETATION

Manual data interpretation is a more traditional and widely used approach. It applies to static data sets that must be exported from a database and processed using third party platforms.

AUTOMATIC DATA INTERPRETATION

Automatic data interpretation involves the processing of real time data within a connected online environment. Correction, harmonisation, cleaning and analytics can all be applied to data as inbuilt functions of a platform. As smart city technologies mature, such automated functions will improve in capability and reliability, providing real time interpretation that can support a variety of high-impact operational outcomes.

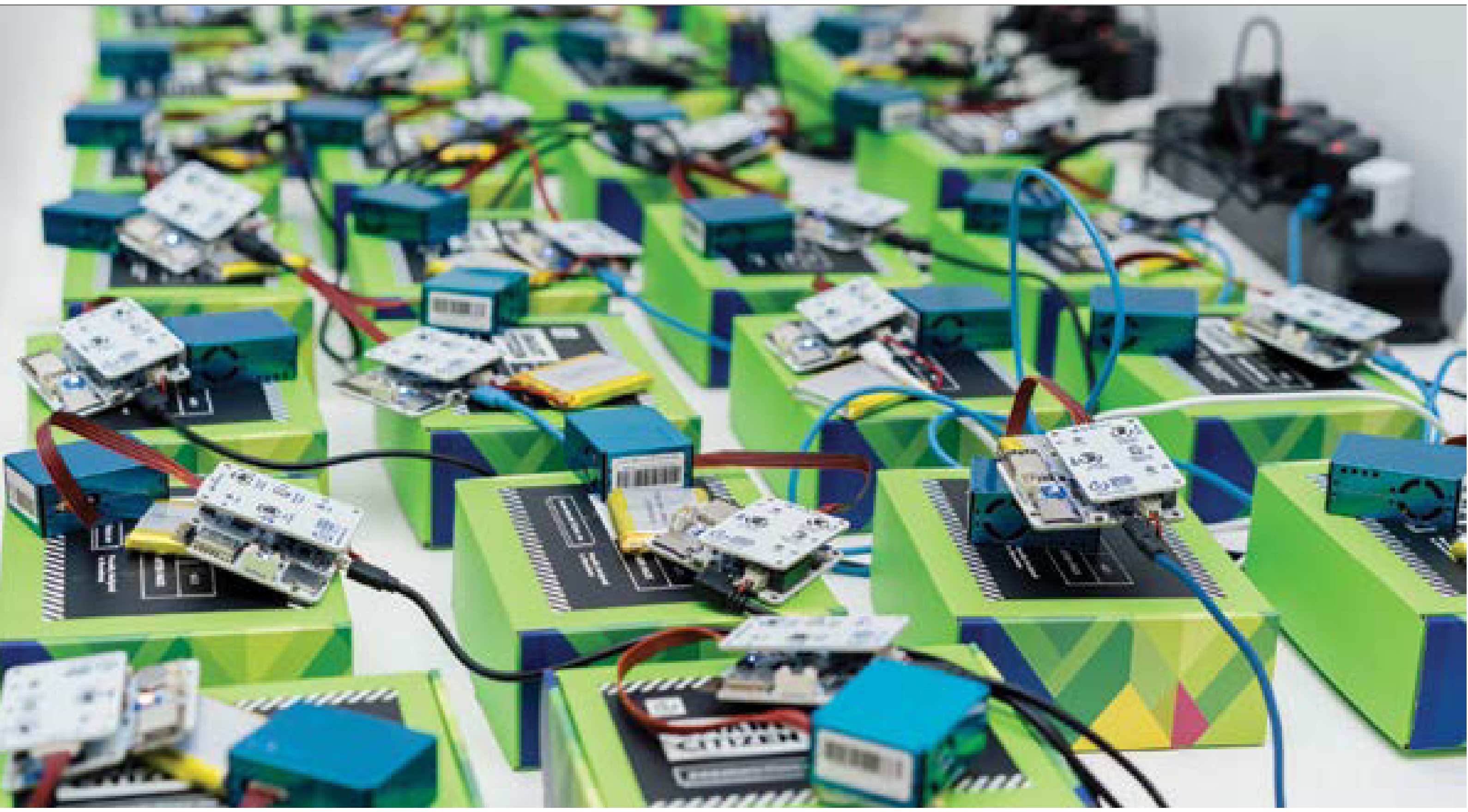


Image: Indoor co-location of Smart Citizen particulate sensors to assess intra-device variability
Source: Creative Commons

1 STEP 1: Data correction and harmonisation

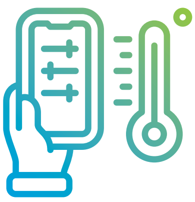
Data correction and harmonisation converts raw data into a useable and readable standard format. This process includes:

CALIBRATION CORRECTIONS

Corrections may be needed to raw data output from devices to address several common issues, including systemic bias for a particular device type; calibration drift over time; inter-device measurement variability; and calibration variation for different types of particulate pollution.

ENVIRONMENTAL INTERFERENCE CORRECTIONS

Sensor response can be significantly influenced by environmental conditions, especially when there are extremes in either relative humidity (RH) or temperature. These are referred to as environmental interference factors. The raw data may need to be corrected to take these factors into account.



DATA HARMONISATION

Data harmonisation involves converting data expressed in different formats into a single, harmonised format (defined by a data schema). Large, harmonised data sets support improved analysis and better project outcomes. Harmonisation is also key to effective data sharing.

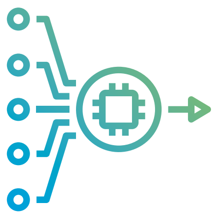


Image source: Creative Commons

2 STEP 2: Data quality control

Data quality control is the process by which interpreted data is checked to ensure it can be trusted and is useable for analysis. This process includes:

FIELD TESTING

An initial device and data verification process, where the 'normal' operation of a device (and the quality of the data that it produces) is verified prior to commencement of your main data collection activities.

DATA CLEANING

Cleaning data means detecting and removing data anomalies, outliers, and other unusable data.

VERIFICATION

Verification of data quality usually happens during standard operations. In this process, data is verified against external references (or through internal cross-verification) to improve user trust in the data and in its ability to support project impacts and goals.

3 STEP 3: Data analysis

Data analysis is the interrogation and utilisation of data to produce insights or functional outputs. These, in turn, support project outcomes and impact creation. Data analysis may include:

STATISTICAL ANALYSIS	Interrogating a set of data to identify patterns and trends, around which insights may be established.
TEMPORAL INTERPOLATION	Inferring values for a fixed point in time, based on data from before and after that point.
SPATIAL AGGREGATION	Combining data from multiple devices in a given area to provide various average values for air quality in that area.
SPATIAL INTERPOLATION	Inferring values for a point on a map – such as a heat map – where there is no data, based on data from other points in the area where there are sensors)
COMPLEX GEOSPATIAL SYSTEM MODELLING	Using various advanced approaches, including heterogenous data synthesis, where data sets with fundamental differences are harmonised; atmospheric pollution dispersion models; and digital twins.
AI AND MACHINE LEARNING APPLICATIONS	Augmenting and enhancing an existing data analytics model, using advanced data analytics capabilities.

Further reading

For further detail and in-depth practical guidance on this topic, please see the OPENAIR Best Practice Guide chapters on data interpretation. These include an overview chapter and chapters on each of the three steps: correction and harmonisation, quality control, and analytics.

FIND OUT MORE AND ACCESS OPENAIR RESOURCES

This factsheet 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. Check the project website for resources and updates on post project collaborations: **www.openair.org.au**

OPENAIR is coordinated by the National Smart Sensing Network (NSSN) and delivered in partnership with the University of Technology Sydney, Australian National University, Western Sydney University.

OPENAIR is made possible by the Smart Places Acceleration Program under NSW Government's Digital Restart Fund.

20231020 F401 Air quality data interpretation Version 2 Final