

Air Quality Data Commons

Abhijit RS, Environmental Defense Fund

Air Sensors International Conference

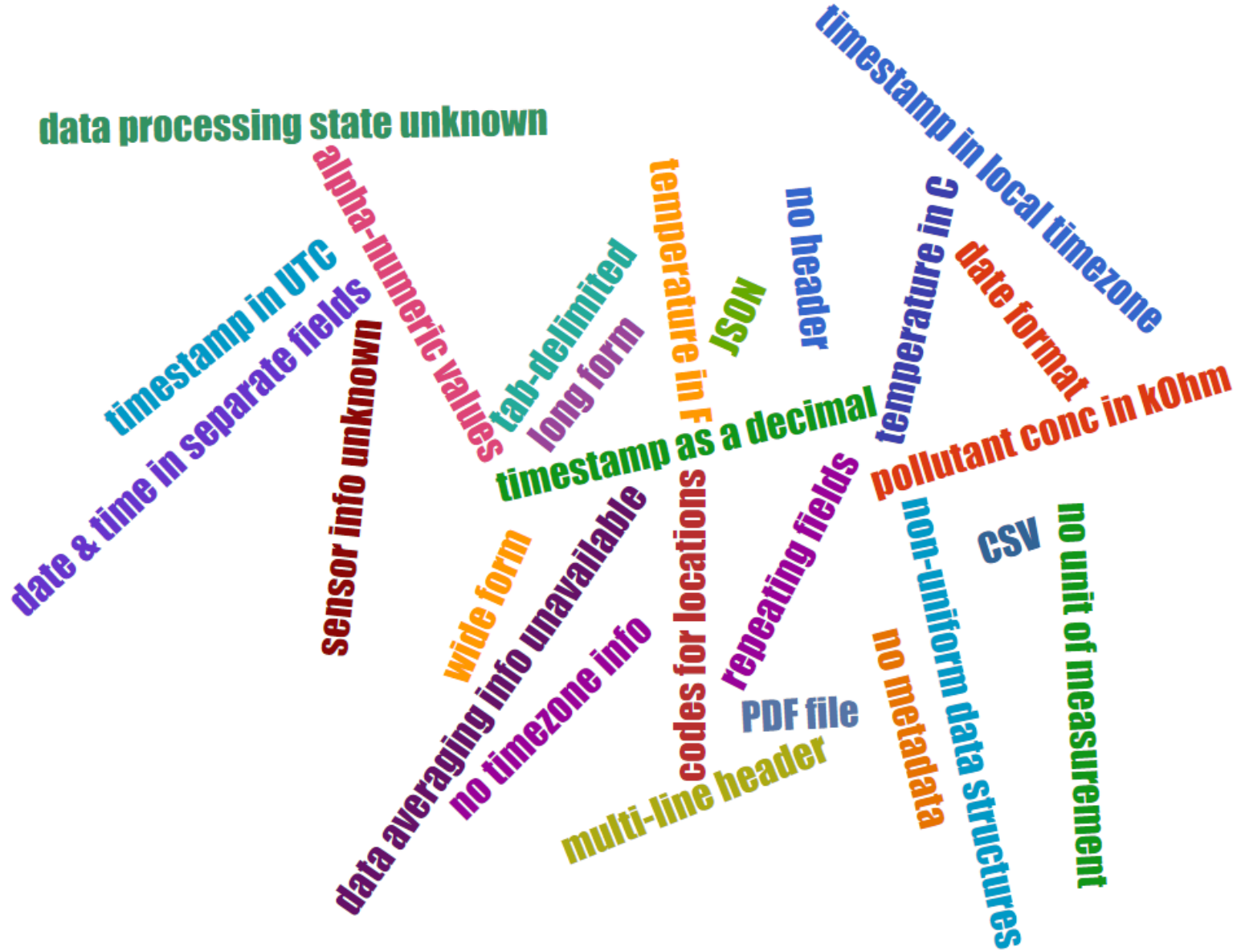
14 Sep 2018



Current Situation



Data Issues



Opportunity



Atmospheric Measurement Techniques

An interactive open access journal of the European Geosciences Union

[EGU.eu](#) | [EGU Journals](#) | [EGU Highlight Articles](#) | [Contact](#) | [Imprint](#)

Volume 6, Issue 10

Atmos. Meas. Tech., 6, 4487–4505, 2015
<http://www.atmos-meas-tech.eu/>
doi:10.5194/amt-6-4487-2015
© Author(s) 2015. This work is distributed under the Creative Commons Attribution 3.0 License.

[Article](#) | [Peer review](#) | [Metrics](#) | [Related articles](#)

Research article

Spatial mapping of ground-based observations of total ozone

K.-L. Chang¹, S. Gillies², and V. E. Fioletov²
¹Department of Statistical Science, University College London, London, UK
²Environment Canada, Toronto, Ontario, Canada

Received: 26 Jan 2015 – Discussion started: 22 Apr 2015
Revised: 29 Sep 2015 – Accepted: 07 Oct 2015 – Published: 23 Oct 2015

Abstract. Total column ozone variations estimated using ground-based stations provide important independent source of information in addition to satellite-based estimates. This estimation has been vigorously challenged by data inhomogeneity in time and by the irregularity of the spatial distribution of stations, as well as by interruptions in observation records. Furthermore, some stations have calibration issues and thus observations may drift. In this paper we compare the spatial interpolation of ozone levels using the novel stochastic partial differential equation (SPDE) approach with the covariance-based kriging. We show how these new spatial predictions are more accurate, less uncertain and more robust. We construct long-term zonal means to investigate the robustness against the absence of measurements at some stations as well as instruments drifts. We conclude that time series analyses can benefit from the SPDE approach compared to the covariance-based kriging when stations are missing, but the positive impact of the technique is less pronounced in the case of drifts.

Citation: Chang, K.-L., Gillies, S., and Fioletov, V. E.: Spatial mapping of ground-based observations of total ozone, Atmos. Meas. Tech., 6, 4487–4505, doi:10.5194/amt-6-4487-2015, 2015.

Air Sensor Workgroup



Careers • Offices • Contact us • For the media



Our work How we get results How you can help About us Blogs

Donate now

Home • Our work • Health • Resources • Air Sensor Workgroup

Air Sensor Workgroup

Collaborative development of standards and infrastructure for air quality data

The Air Sensor Workgroup (ASW) is a broad-based collaboration that support the rapidly growing community of people developing and using sensor-based air quality devices. The ASW's goal is to advance this technology sector with infrastructural support including adoption of data standards and provision of an open data platform, the Air Quality Data Commons (AQDC).

EDF brought this group together and has been leading the effort to achieve democratization of air quality data to improve the understanding of people's exposure to air pollution. This group furthers the principles of air quality data being open and FAIR (Findable, Accessible, Interoperable, and Reusable).

Data standards will enable the use of standardized terms and data structures, identify key data elements necessary to describe the data, and provide common data exchange formats.

The AQDC will provide for centralized data storage and enable analysis and visualization of neighborhood-level air quality data. The AQDC will also provide infrastructure to perform big data computation and allow for data products to understand the sensor response and encourage the community to develop data-driven solutions to cleaner air.

ASW resources

Date and timestamp guidelines
Derived from standards of ISO, IETF, and W3C

Standardized terms

External resources

- AirNow
- AQ-SPEC sensor evaluations
- Marine Metadata Interoperability Ontology
- AMS Glossary

Our experts



Abhijit RS
Data Architect
Email



Millie Chu Baird
Associate Vice President,
Office of the Chief
Scientist

Date & timestamp guidelines

By Air Sensor Workgroup

Status of this document

The Air Sensor Workgroup (ASW) adopted the date and timestamp guidelines on 17 January 2017.

Abstract

This document defines the Date and Timestamp Guidelines for use in the field of air quality measurement and monitoring. It has been derived from [ISO 8601](#) standard, IETF [RFC 3339](#) and the [W3C profile](#).

Users and applicability

ASW strongly encourages the manufacturers of low cost sensors, researchers working on air quality and any other air quality data generators and users to use these guidelines.

Purpose of guidelines

Numerous individuals and organizations across the globe have spent effort to measure air quality data and determine the impact of air pollution on human health. However, most of them have been isolated efforts. ASW sees tremendous value in sharing data across data owners so that researchers and other interested parties can take advantage of the vast amount of data to create air quality data products that can help communities worldwide. To this effect, ASW has been developing standards for data generation, storage and exchange.

The Date and Timestamp guidelines apply to data generation and storage by the sensors and their backend database systems, thus facilitating accurate, reliable and efficient exchange of data across various data owners and data users.

Format for data generation

Use Epoch time (aka Unix time, [POSIX time](#)) which is time in seconds since Unix Epoch (1970-01-01T00:00:00Z) as a 64-bit unsigned integer at the point of generating date/time value by the device.

- It simplifies date arithmetic
- Use 64-bit integer as the data type so as to avoid data overflow in the year 2038
- Use microseconds granularity in order to support measurements at frequencies higher than once per second (1Hz)
- Use this format for transmitting data from the sensor to the backend server
- For writing to the logs, including logs stored on the sensor or the sensor system, use the human readable format and associated guidelines described in the next section. This increases the ease and efficiency of onsite troubleshooting and maintenance.
- Tools and libraries to convert to human-readable formats are available and can be applied just prior to visualization

Key considerations

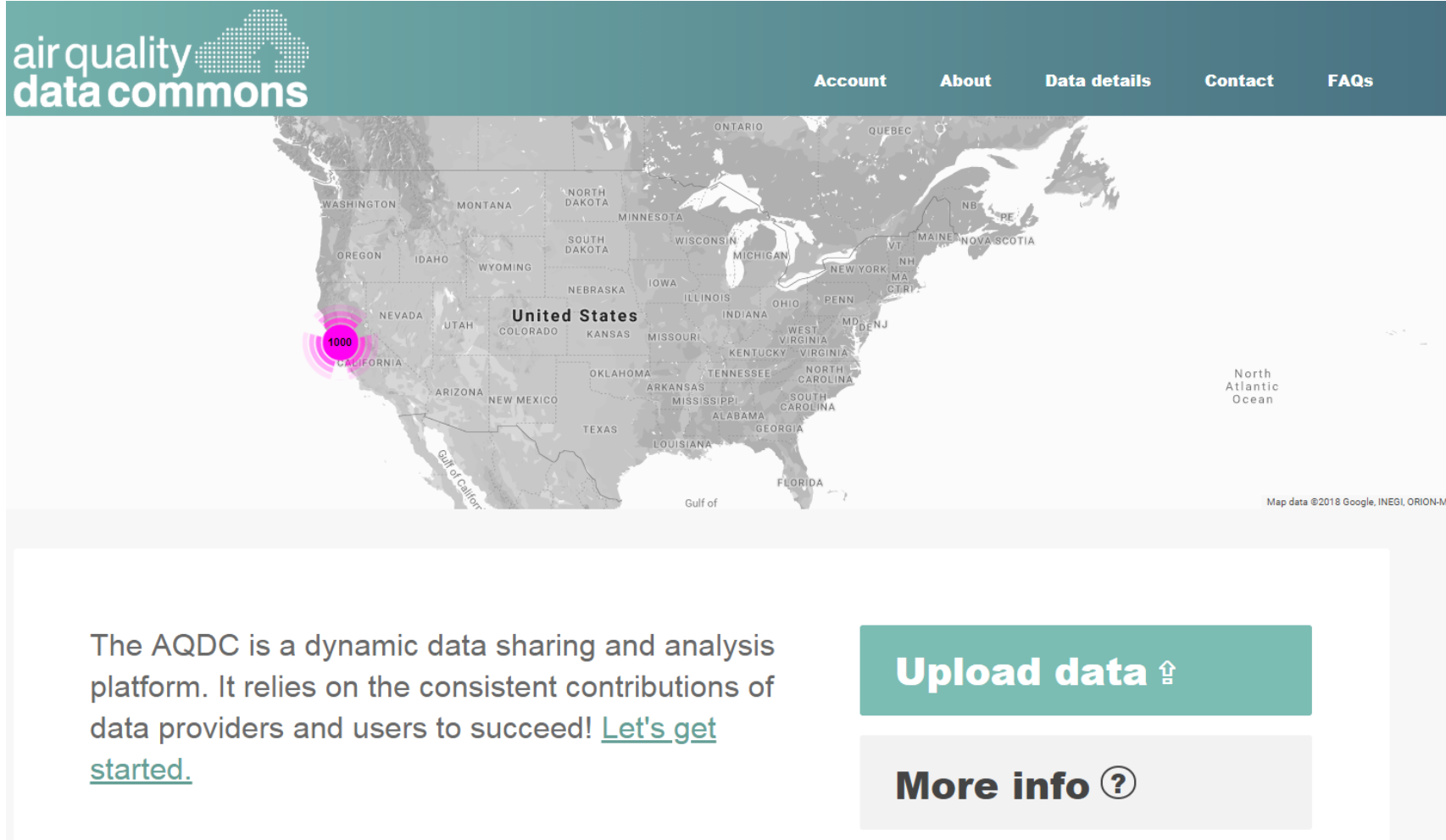
- Multiple application tiers on the device such as I2C/IC/RTC, microprocessor and software may generate the timestamp value. All these tiers must support 64-bit integer. If any of these support 32-bit only, then the Epoch time will overflow on 19-Jan-2038 and reset to 13-Dec-1901.
- Some RTCs have a ceiling on the year they support like 2099, 2100, etc. Pay attention to what is supported and have a plan for subsequent time period.
- Use appropriate data type in the software programs to support 64-bit integer values
- If there is a need for sub-second measurements, explore the built-in support provided by the databases and programming languages (see Appendix).

Format for data storage

The requirements for data storage are driven not only by storage optimization and retrieval (I/O) performance considerations but also the subsequent data usage, including the visualization layer. Store the timestamp value as Epoch time (as received from the

<https://www.edf.org/asw>

Air Quality Data Commons (AQDC)



The screenshot shows the AQDC website interface. At the top left is the logo "airquality data commons" with a stylized map icon. To the right of the logo are navigation links: "Account", "About", "Data details", "Contact", and "FAQs". Below the navigation is a map of the United States with a pink circular marker over California labeled "1000". Below the map is a text box with the following content:


The AQDC is a dynamic data sharing and analysis platform. It relies on the consistent contributions of data providers and users to succeed! [Let's get started.](#)

Below the text box are two buttons: "Upload data" with an upward arrow icon, and "More info" with a question mark icon.



<https://aqdatacommons.org>

Air Quality Data Commons (AQDC)

- ✓ Centralized platform for AQ data storage
 - ✓ Standardized data structure
 - ✓ Compute infrastructure for data analysis
 - ✓ Open access data, open source software
- 

Air Quality Data Commons

Data Sources

- Low- & medium-cost AQ sensors
- Stationary and mobile
- Global geographic scope
- Integration of regulatory data (forthcoming)

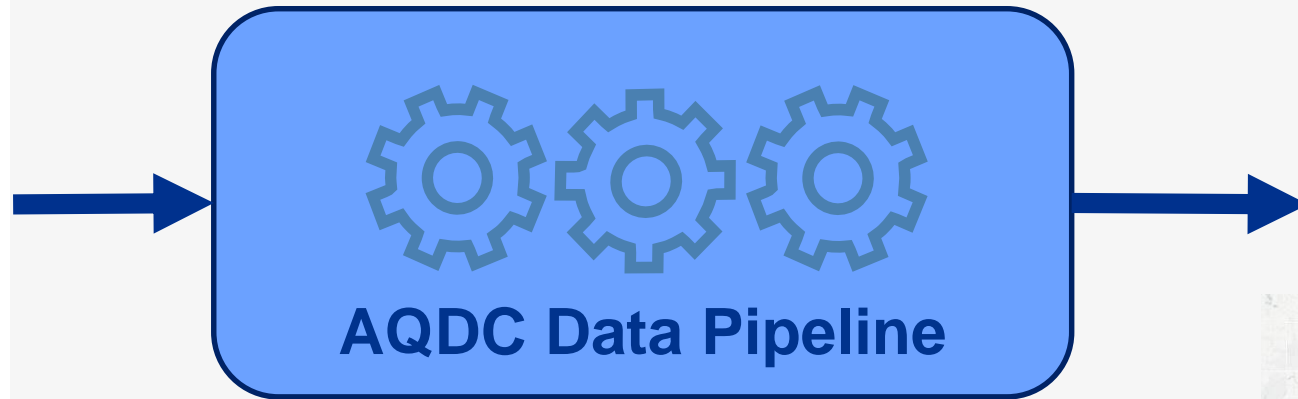
Data Users

- Air Quality specialists & enthusiasts
- Researchers
- Academics
- Sensor Manufacturers
- Government Agencies
- Citizen Scientists
- Open source software contributors

Air Quality Data Commons (AQDC)

The screenshot shows the 'air quality data commons' website with a navigation bar containing 'Account', 'About', 'Data details', 'Contact', and 'FAQs'. The main heading is 'Upload data'. Below it, a sub-heading reads: 'You can upload data below by selecting a file, and then completing some configuration settings. Once ready, you can submit the data with the settings for processing.' There are two main sections: '1 Select file' and '2 Set general parameters'. The '1 Select file' section includes instructions: 'CSV format only' and 'Column headers required', and a 'Choose File' button with the text 'No file chosen'. The '2 Set general parameters' section contains several input fields, each with a red asterisk indicating it is required: 'Data collection project/campaign name', 'Name of sensor deployment site', 'Sensor system deployment date', 'Data license, as determined by data owner', 'Effective date for this metadata', 'Data upload version', and 'Timestamp format'. Each field has a corresponding input box or dropdown menu.

Upload your data

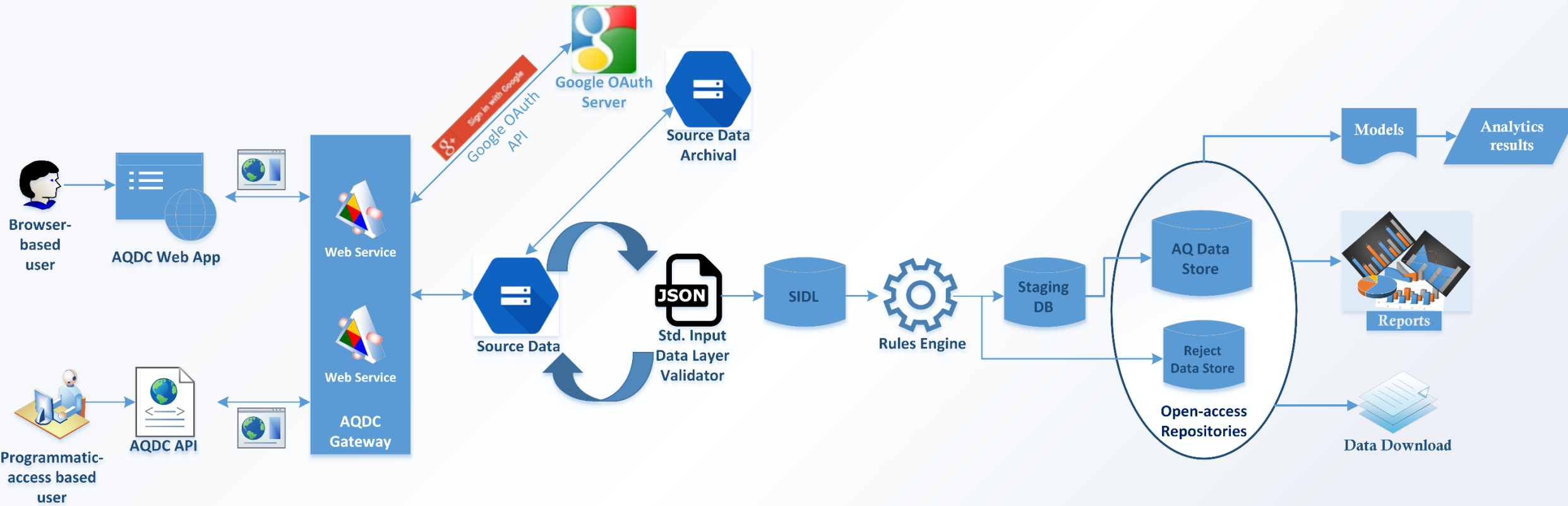


AQDC processes your data



Visualize your data

AQDC Application Architecture



Air Quality Data Commons (AQDC) Application Architecture

AQDC Features

- Open access data repository, open source software
- Cloud-native platform to store, analyze & collaborate on air quality data
- Data standardization
 - Date & timestamp guidelines
 - *Standardized Input Data Layer (SIDL)
- API and Browser-based access
- *Data quality – Rules Engine
- *Reject Data Store
- *Metadata repository
 - Sensor data sheets
- Visualization - out of the box & self-serve
- *Data exploration & analysis tools
- Widgets for frequently performed analyses (forthcoming)



Thank You

Abhijit RS
arudrapatna@edf.org
415-293-6105

