

# Accessing AQS Data

Chris Chapman, US EPA

National Ambient Air Monitoring Conference

August 10, 2016 – St. Louis, Missouri





# Outline

- Background
- Trends in Use
  - Popularity of data
  - Who's getting the data
  - What are they getting
- Data Sources
  - Air Data
  - AQS API
  - AQS Reports



# Background

- Several years ago, OAQPS embarked on an effort to reduce the number of places public could go to find data
- Pick certain ones to enhance, others to decommission, put all that remain in a common place
- The new umbrella site = AirData
- AQS specific
  - Rebranding and rearchitecture of AirData (now AQS API)
  - Added addition AMP series reports, updated Existing reports (ex: Design Values report for 2015 Ozone NAAQS, added new QA reports (AMP600, AMP251, and AMP256))



# Sites

- Air Quality data is in high demand
- AirData:
  - <https://www3.epa.gov/airquality/airdata/>
  - Reports, graphs, and downloads of criteria data
  - For analysts and informed public
- AQS API:
  - <https://aqs.epa.gov/api>
  - Web service for querying any raw data
  - For analysts and scientists
- File Downloads:
  - [http://aqedr1.epa.gov/aqsweb/aqstmp/airdata/download\\_files.html](http://aqedr1.epa.gov/aqsweb/aqstmp/airdata/download_files.html)
  - Bulk downloads of most requested raw data
  - For analysts and scientists

## Query AirData

Download data for a specific location and time for any pollutant by interactively querying the AQS Data Mart database.

A user ID and password are required to use this page. (Register for an account.)

More information is available on how to use this interface. Please see the full AQS documentation (OASIS version) for data services available, REST service construction, output formats, etc.



# Trends in Use

- Air Quality data is in high demand
- AirData: ~7,000 queries per month
- AQS API: ~7,000 queries per month
- File Downloads: ~45,000 per month
- 25 billion rows of data queried since Jan 01, 2009
  - 850 million since Jan 01, 2016
- 500 billion rows of data downloaded in files since Jan 01, 2016
- Every value input into AQS today will be viewed/used thousands of times over the next few years



# User Categories

- Heuristics based on contacts and email addresses
  - Academics – many(!) public health / epidemiologist users
  - NGOs
    - Mostly health related
  - Corporations
    - Regulated
    - App developers
      - Validation of small “sensors” (public/crowdsourced networks)
    - Business reasons: battery life, corrosion, etc.
  - SLTs (air quality and health departments)
  - Academics - model / satellite verification
  - Academics - Big Data / Data Science mining
  - Interested public



## Types of Data (Anecdotal Trends)

- PM2.5 & gaseous criteria pollutants continue to be very popular
- Met data amazingly popular
- PM2.5 speciation interest continuing to increase
- Continuing interest in HAPs/VOCs



# Data Sources

- Air Data
- AQS API
- AQS - Reports





# What is Air Data?

- Publicly available tool...no registration required
- Good for monitor network information, AQI (Air Quality Index) information and monitor, county and CBSA level summary information
- Can download raw data via web service to AQS AQI
- Access at <https://www.epa.gov/airdata/>
- RSS Feed at <https://www.epa.gov/airdata/rssairdata.xml>
- Information at [https://www.epa.gov/airdata/ad\\_basic.html](https://www.epa.gov/airdata/ad_basic.html)



# Air Data Areas

- AirData has many “areas”
  - Each targeted towards a different type of user
    - Reports
      - Tabular data (annual) that can also be downloaded
      - Criteria pollutants only, form of the standard
    - Visualizations (Plots)
      - Graphical data (daily data based) that can also be downloaded
      - Criteria pollutants only, form of the standard
    - Maps
      - Geographical data that can also be downloaded
      - Criteria, Speciation, and HAPS



# Air Data Areas (Cont'd)

## Download

- Daily (criteria pollutants only, form of the standard)
- Raw (all pollutants) “AQS API”
  - Registration required (3,200 registered users)
  - REST web services – use as GUI or API
- Files
  - Pre-generated files of most commonly requested data
  - Annual, daily, and hourly
  - National, by year back to 1990 (~800 files)
  - Updated twice a year, automated, only files that change
  - CSV (not AQS) format
  - Replaces TTN data download page

# AirData Areas



The screenshot shows the EPA AirData website interface. At the top, there is a navigation bar with links for 'LEARN THE ISSUES', 'SCIENCE & TECHNOLOGY', 'LAWS & REGULATIONS', and 'ABOUT EPA'. Below this is the 'AirData' header with the tagline 'Access to monitoring data'. The main content area is divided into several sections:

- Basic Information:** A section for general information.
- Visualize Data:** A section with four options: 'AQI Plot' (Compare AQI values for a location & time), 'Tile Plot' (Plot daily AQI values for a location & time), 'Concentration Plot' (Generate a time series plot for a location), and 'Concentration Map' (Generate concentration maps for a specific time). There is also a 'More Tools' link.
- Download Data:** A section with a 'Download Data' button and text: 'Download daily or raw data to a spreadsheet or file.'
- Interactive Map:** A section with a map of the United States showing monitor locations and text: 'See where air quality monitors are located, get information about the monitors, and download monitor data. Select which monitoring networks to display.'
- Reports:** A section with four report types: 'Air Quality Index' (Display an annual summary of AQI values), 'Air Quality Statistics' (Display standards-related summary data by city or county), 'Monitor Values' (Display criteria pollutant summary data), and 'Monitor Values - HAPs' (Display Hazardous Air Pollutant (HAP) summary data).
- Latest News:** A section with two news items: 'April 2, 2014 - A Monitor Report for Hazardous Pollutants is now available.' and 'March 4, 2014 - Pre-generated files of data for the entire nation are now available at Download Data Files. These are files of our most commonly requested parameters at the annual, daily, and hourly summary levels.'
- Sources of Air Data:** A section with a list of sources: 'AirNow - forecasts and real-time data', 'AirCompare - help with planning a move or vacation', 'Air Emission Sources - national, state, and county summaries', and 'More Sources'.

Four red callout boxes highlight specific features: 'REPORTS' (pointing to the Reports section), 'VISUALIZE' (pointing to the Visualize Data section), 'DOWNLOAD' (pointing to the Download Data section), and 'MAP' (pointing to the Interactive Map section). A fifth red callout box is empty and points to the Sources of Air Data section.

At the bottom of the page, there is a footer with links for 'EPA Home', 'Privacy and Security Notice', and 'Contact Us'. It also includes icons for 'News by E-mail', 'EPA Mobile', 'Widgets', 'News Feeds', and 'Podcasts', along with the EPA logo.



# What is the AQS Data Mart / AQS API?

- Data Mart is a database containing all of the data from AQS
  - Powers AirData and AQS API, and has same architecture as AQS.
  - AQS Data Mart also includes ozone data for AQS monitors that are reported to AirNow
- Any “application” can connect to it (within tech/security limits)
  - AirData, KML files, Ozone Watch, AirCompare, Data.gov
- Intended users:
  - Air quality data analysts in the regulatory, academic, and health research communities
  - Those who need to download large volumes of detailed technical data
  - People who need the AQS data for the application they are building



# AQS API INTERFACE (aqs.epa.gov/api)

## Query AirData

Download data for a specific location and time for any pollutant by interactively querying the AQS Data Mart database.

A user ID and password are required to use this page. ([Register for an account.](#))

More information is available on [how to use this interface.](#)

Please see the full [API documentation \(RAML Version\)](#) for data services available, REST service construction, output formats, etc.

User Name:	<input type="text"/>
Password:	<input type="password"/>
Query Type:	<input type="text"/>
Output Format:	<input type="text"/>
Parameter Class:	<input type="text"/>
Parameter Code:	<input type="text"/>
Begin Date:	<input type="text"/> (yyyymmdd)
End Date:	<input type="text"/> (yyyymmdd)
Min. Latitude:	<input type="text"/>
Max. Latitude:	<input type="text"/>
Min. Longitude:	<input type="text"/>
Max. Longitude:	<input type="text"/>
State Code:	<input type="text"/>
County Code:	<input type="text"/>
Site:	<input type="text"/>
CBSA:	<input type="text"/>
CSA:	<input type="text"/>
Duration:	<input type="text"/>
FRM Only:	<input type="checkbox"/>
<input type="button" value="Submit"/>	



## AQS Reports (AMP xxx)

- Briefly mentioned earlier
- Trend also continues upward
  - 71,000 in CY 2015, 39,000 is 1H CY2016. This averages to more than one every 2 minutes during business hours
    - Does not include Load Summary/Error reports
    - Does include automatically generated SG inventory report
- We'd like to retire some, but all are used
- We've added more, hopefully helpful
- We're working on streamlining



# AQS Reports Interface (R31)

Standard Report Criteria Selection (National Air Data Group)

Criteria Set | Data Selection | Sort Order | Report Options | Retrieve Reports

Criteria Set  Desc

Owner   Type

Report Code  Report Name

Run Online  Send via Email

**Report Outputs**

**Generate Report**





# Reports by Category

Report Category	Reports
Site and Monitor Metadata	AMP380 (Site Description), AMP390 (Monitor Description), AMP220 (Monitor Networks)
Detail Data Reports	AMP500(Extract Site/Monitor Data), AMP 501 (Extract Raw Data), AMP503 (Extract Blanks Data), AMP350, AMP350MX, AMP350P, and AMP350NW (Raw Data)
Summary Data Reports	AMP500(Extract Site/Monitor Data), AMP 501 (Extract Raw Data), AMP503 (Extract Blanks Data), AMP350, AMP350MX, AMP350P, and AMP350NW (Raw Data)
Certification and QA	AMP600 (Certification Evaluation and Concurrence),AMP256 (QA Data Quality Indicator),AMP251 (QA Raw Assessment) AMP504 (QA Data Extraction)
Others	AMP480 (Design Value), AMP360 (Raw Data Qualifiers), Reports not run from R31 interface (AMP395, AMP396, Load, Scan, Post).



# AQS Reports Documentation

- User's Guide: <https://www.epa.gov/aqs/aqs-users-guide-0>
- Data Retrieval Guide currently being updated. When complete will have full explanation of all reports

# Intended audience

AQS API	AQS	Air Data
<ul style="list-style-type: none"><li>• Air quality data analysts in the regulatory, academic, and health research communities</li><li>• People familiar with the data</li><li>• An AQS account is not needed</li></ul> <p>8/15/2016</p>	<ul style="list-style-type: none"><li>• Those who provide data from their federal, state, local, or tribal agency to the EPA</li><li>• Standard reports are available for QA</li></ul> <p>U.S. Environmental Protection Agency</p>	<ul style="list-style-type: none"><li>• Designed for public access (no registration required)</li><li>• Also useful for researchers and analysts who can download raw data</li></ul> <p>19</p>

# Registration

AQS API	AQS	Air Data
<ul style="list-style-type: none"><li>• Requires password</li><li>• Self-registration at home page</li><li>• Anyone can register</li></ul>	<ul style="list-style-type: none"><li>• Registration at AQS site: <a href="https://www.epa.gov/aqs/aqs-user-registration">https://www.epa.gov/aqs/aqs-user-registration</a></li></ul> <p>Must go through S/L/T/Regional approval process</p> <ul style="list-style-type: none"><li>• Must be Fed or State/Local/Tribal employee or designated contractor</li></ul>	<ul style="list-style-type: none"><li>• Not required</li></ul>

# Examples of output

AQS API	AQS Reports	Air Data
<ul style="list-style-type: none"><li>• CSV</li><li>• Pipe delimited</li><li>• JSON</li></ul>	<ul style="list-style-type: none"><li>• Acrobat (pdf)</li><li>• Workfile (txt)</li></ul>	<ul style="list-style-type: none"><li>• On screen</li><li>• Downloadable as csv</li><li>• Visualizations (jpg, gif)</li></ul>

# Update frequency

AQS API	AQS	Air Data
<ul style="list-style-type: none"><li>• Weekly</li><li>• Sunday night</li><li>• Updated by Automated Process</li></ul>	<ul style="list-style-type: none"><li>• Continuously Updated by State / Local / Tribal Organizations</li><li>• Summaries Re-Computed when Posted</li></ul>	<ul style="list-style-type: none"><li>• Same schedule as AQS API</li></ul>



## A moment of silence for Discoverer

- An Oracle tool that is no longer supported by Oracle although it is still available to AQS users
- May be decommissioned by NCC at any time
- Replacement still under consideration / prioritization



## Conclusion

- Demand for data increasing
- Keep up the good work – more and more people using and depending on the data
- We're continually adapting to new requirements and new constraints (technology, budget, time, etc.)
- Suggestions welcome