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://aqsdr1.epa.gov/aqsweb/aqstmp/airdata/download_files.html
  - Bulk downloads of most requested raw data
  - For analysts and scientists

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

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
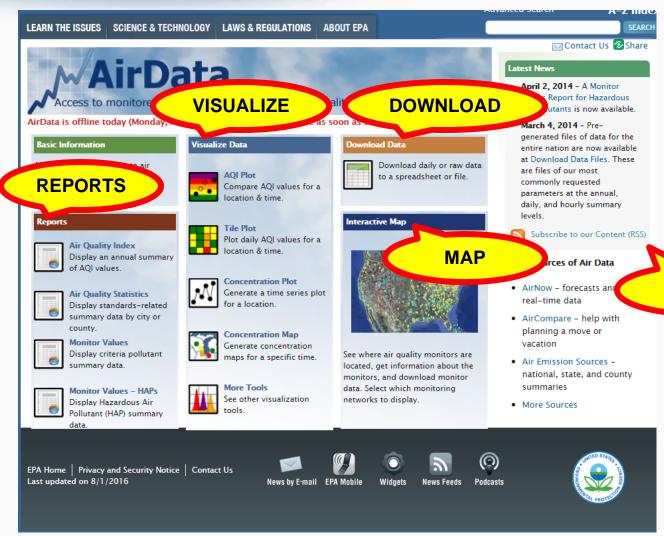   - Criteria, Speciation, and HAPS

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

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

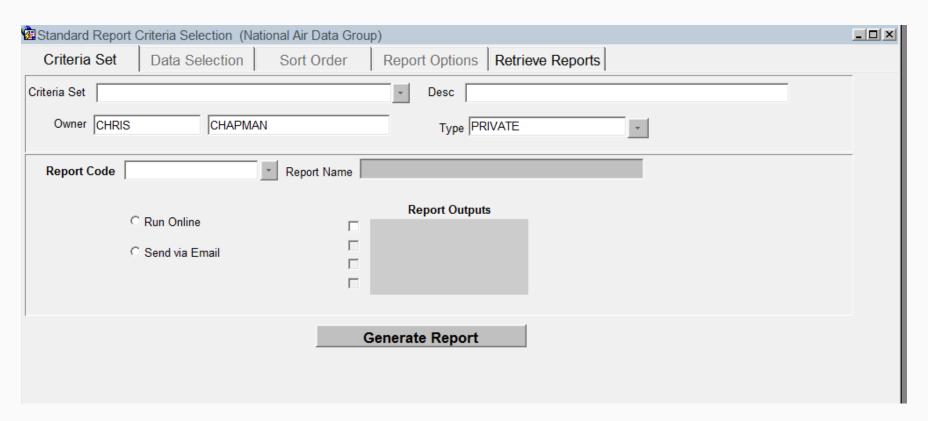# AQS API INTERFACE (aqs.epa.gov/api)

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)

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
• Air quality data analysts in the regulatory, academic, and health research communities • People familiar with the data • An AQS account is not needed	• Those who provide data from their federal, state, local, or tribal agency to the EPA • Standard reports are available for QA	• Designed for public access (no registration required) • Also useful for researchers and analysts who can download raw data

U.S. Environmental Protection Agency

Registration

AQS API	AQS	Air Data
• Requires password • Self-registration at home page • Anyone can register	• Registration at AQS site: https://www.epa.gov/aqs/aqs-user-registration Must go through S/L/T/Regional approval process • Must be Fed or State/Local/Tribal employee or designated contractor	• Not required

Examples of output

AQS API	AQS Reports	Air Data
• CSV • Pipe delimited • JSON	• Acrobat (pdf) • Workfile (txt)	• On screen • Downloadable as csv • Visualizations (jpg, gif)

Update frequency

AQS API	AQS	Air Data
• Weekly • Sunday night • Updated by Automated Process	• Continuously Updated by State / Local / Tribal Organizations •Summaries Re-Computed when Posted	• Same schedule as AQS API

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