Obtaining AQS Data

Discoverer Replacement
The Replacement plans for Discoverer are as follows:

- The AQS Federal Team will construct a set of “views” that correspond to each type of data in AQS, i.e. Sites and monitors, raw data, summary data, and QA data.
  - Each view will contain **all** of the site and monitor metadata pertinent to the information provided by the view.
  - Each summary view will be targeted to a specific Pollutant Standard and Exceptional data type.
- A web application will be developed (under AirData) to allow users to query data from these views; it will output CSV files.
- Federal users (inside the EPA firewall) will be able to use off-the-shelf tools, such as TOAD or SAS, to directly access these views.
- Third-party tools, such as SAS or Excel, will be used for analysis.
- The views will be enhanced as needed in response to user requests.
Alternatives Considered

- Oracle Business Intelligence Tool: Cost Prohibitive (NCC chargeback ~$160k/yr). Other commercial BI tools comparable in price.
- GUI Query Builder Tool (Like discoverer interface): All of the open-source tools that we found were client based (i.e. run directly on windows) and EPA security staff do not allow SQL-net through EPA firewall.
Rationale for Plan

• Generation of CSV Files: Most analysis tools can easily read CSV.
• Discoverer provides a subset of the graphics and analysis capabilities of current spreadsheet products (Excel or Open Office).
• More than 90% of user support for discoverer related to the following two issues:
  – Joining multiple tables together to generate the desired output.
  – Selecting too many records because of Exceptional Data Types and Pollutant Standards
Categories of Views

- Sites, Monitors, and Samplers
- Detailed Data (e.g. Raw Data, Blanks, & QA)
- Generic/Non-regulatory Summaries
- Regulatory Summaries
- Reference Data (AQS Codes)?
  - Are these necessary? The other views will include code descriptions wherever code values are used.
Site-Monitor Views

- Views will be created for Sites, Monitors, and Samplers.
- The Site view will have all of the site identifying information and metadata
- The Monitor view will have all of the site metadata and monitor metadata
- The Sampler view will have all of the site metadata and sampler metadata
- There are open issues as to handle date-range information. Initial plan is to include present state of site or monitor; i.e. the PQAO in the monitors view will be the current PQAO.
- The Web user interface will allow queries by Network.
Detailed Data Views

- These views will contain the un-summarized data submitted to AQS:
  - Raw Data
  - Blanks Data
  - QA Data
    - 1-Point_QC
    - Annual_PE
    - Flow Rate Verification
    - Semi-Annual Flow Rate Audits
    - And etc.
Generic Summaries

• These views will provide summaries of data, ignoring Exceptional Event Flagging; each will contain all site and monitor metadata.
• Multi-Hour Summaries
  – 8-Hour Ozone
  – 8-Hour CO
  – 3-Hour SO2
• Daily summaries
  – Daily_Observed_Data (will contain all observed durations)
  – Daily_Ozone_8_hour
  – Daily_CO_8_hour
  – Daily_SO2_3_hour
• Lead Monthly Summaries
• Lead 3-month rolling averages
• Quarterly Summaries (for all above durations)
• Annual Summaries (for all above durations)
These views will contain summaries used for regulatory purposes (i.e. NAAQS Determinations); i.e. they will be pollutant standard specific and exclude concurred EE flagged values.

- Multi-Hour Summaries
- Monitor Daily Summaries
- Monitor Quarterly Summaries
- Monitor Annual Summaries
- Site Daily Summaries
- Site Lead Monthly Summaries
- Site Lead 3-Month Rolling Averages
- Site Quarterly Summaries
- Site Annual Summaries
- Design Value Views
Multi-Hour Summaries

- The multi-hour summaries will contain the rolling or block averages specified by 40 CFR Part 50 for specific pollutants. These include:
  - Ozone 8-Hour 2006 Standard
  - Ozone 8-Hour 1997 Standard
  - CO 8-Hour 1971 Standard
  - 3-Hour SO2 1971 Standard
  - Note: 24-hour PM2.5 and SO2 block averages not included since these exactly coincide with a day and are redundant.
Monitor Summaries Daily, Quarterly, Annual

- Ozone_1_hr_1997
- Ozone_8_hr_1997
- Ozone_8_hr_2006
- Lead_2009
- CO_1_hr_1971
- CO_8_hr_1971
- SO2_1_hr_2010
- SO2_3_hr_1971
- SO2_3_hr_1971
- SO2_24_hr_1971
- SO2_Annual_1971
- NO2_1_hr_2010
- NO2_Annual_1971
- PM25_24_hr_2006
- PM25_Annual_2006
- PM25_24_hr_2013
- PM25_Annual_2013
- PM10_24_hr_2006
Site Combined Summaries

- Site Daily
- Site Lead Monthly
- Site Lead 3-Month Rolling
- Site Quarterly (excludes lead)
- Site Annual (excludes lead)

Pollutants:
- Lead_2009
- NO2_1 HR_2010
- PM25_24 hr_2006
- PM25_Annual_2006
- PM25_24 hr_2013
- PM25_Annual_2013

Note: PM10 expected soon; may generalize to all criteria pollutants
Design Values

• Views for where 3-year design value specified by 40 CFR Part 50
  – Ozone_8-hour_2006
  – Ozone_8-hour_1997
  – PM25_2013
  – PM25_2006
  – SO2_2010
  – NO2_2010
  – Lead_2009
  – PM10_2006