Presentation of the AES*Online and AES*XML Emission Inventory Application

Bruce E. Lane, Venkat Ramachandran, Randy Lettich, Kris Minnich, Arjuna Sarode, Barb Garofalo
Ciber, Inc., 650 Wilson Lane, Suite 200, Mechanicsburg, PA 17055
blane@ciber.com

ABSTRACT
The Pennsylvania Department of Environmental Protection Bureau of Air Quality (PADEP) commissioned Ciber, Inc to create a full-featured, web-based system for the collection, reporting, review, and management of emission inventory data from regulated industries within the Commonwealth. To fulfill this initiative, Ciber, Inc. designed, developed, and implemented the *AES*Online/AES*XML* suite of custom applications. *AES*Online* is a comprehensive solution comprising two secure online applications: one that provides facilities with a mechanism to enter, edit, and submit Annual Emission Statement (AES) data to the PADEP and an associated application that allows PADEP staff to review, reject/accept, and manage the AES data. *AES*XML* applies the Commonwealth's AES XML schema to provide facilities with the ability to upload and submit AES data directly from their own databases or spreadsheets. The *AES*Online* and *AES*XML* systems streamline the reporting of emissions inventory data, offer real-time data validation and data quality feedback, and automate the emission inventory reporting and PADEP data review processes.

The initial implementation of *AES*Online* in January 2004 was met with a high level of interest and participation in the regulated community and significantly decreased the level of effort required for AES processing in the 2004 reporting season. The applications were enhanced and optimized to further strengthen the utility and popularity of the system, and participation levels increased for 2005. A pilot of the fully automated reporting option *AES*XML* was implemented in January 2005, completing the suite of PADEP online reporting options. AES*XML shares many of the key design features of *AES*Online* and is focused on more fully automating the reporting process for the regulated community.

The *AES*Online* and *AES*XML* systems run in an Oracle environment using the Oracle 10g Application Server with access to back-end Oracle 9i databases. The Java-based user interfaces use the J2EE architecture. Oracle's suite of XML tools is employed for XML parsing and management. The development effort incorporated the use of several tools and techniques, including Java, JSP, XML, XSL, HTML, PDF and PL/SQL.

These innovative applications provide PADEP with an integrated solution to efficiently and effectively collect, query, and report air emission inventory data. *AES*Online* and *AES*XML* will serve as a model for future automation initiatives in the emission inventory arena.

INTRODUCTION
The Federal mandates for environmental data collection and reporting are ever-expanding and increasingly complex. Changes in government administrations, heightened public concern, and tighter fiscal economies have rendered this a daunting task for both public and private entities. In anticipation of these increased requirements, the Pennsylvania Department of Environmental Protection (PADEP) Bureau of Air Quality (BAQ) embraced an approach over the last five years that included structured planning and analysis, involvement in State and Federal improvement initiatives, and the implementation of automated solutions.

BAQ invested its resources in strengthening and improving their air emission inventory processes. In 2002, BAQ commissioned CIBER, Inc. to design, develop, and implement web-bases solutions to:

- Streamline the emission inventory process
- Increase data quality, access, and usability
- Save costly preparation and processing time
- Cut recurring financial and resource costs
- Ease reporting burden on Industry
- Refocus staff (field work vs. pushing paper)
- Reduce paper use, postage, and filing space

CIBER, Inc. worked with BAQ to create *AES*Online* and *AES*XML*, a suite of full-featured web applications designed specifically for the collection of emission inventory data from industries within the Commonwealth, the review of the data by AQ staff, automated communication of data issues and their correction, and automatic loading of the finalized data into the Bureau's Air Information Management System (AIMS) database. These two systems, working in conjunction with AIMS, met the Bureau's goals primarily by:

- Automating the historical paper-driven data collection and review process
- Removing the errors caused by multiple data entry steps
- Instituting real-time data validation and quality assurance functions
- Reducing the need for large-scale report generation and distribution of hardcopies to Industry
- Reducing paper handling and data entry by AQ staff
- Creating automated tools and repeatable processes
- Creating higher quality data that is available for quicker turnaround and use
- Decreasing the time required for emission reporting and review processes

This paper describes many of the technical aspects of the *AES*Online* and *AES*XML* applications. It presents some of the history and background of the systems and the Bureau's approach to making them a reality. The main focus is to describe the high-level features and functions of the applications, including the general process, common features, and web pages and features unique to each. The intent of the paper is to provide other environmental data management organizations with an appreciation for the completeness and maturity of the systems created by the Pennsylvania Bureau of Air Quality.

BODY

Pennsylvania's path to automated emission inventory management has a long and purpose-driven history. The Bureau has been active in various automation initiatives within its organization and with outside organizations. Beginning in the mid-1990s, the Bureau became involved with the United States Environmental Protection Agency (EPA) and other State and local management agencies through the Emission Inventory Improvement Program (EIIP) to create the data model that would evolve into today's National Emission Inventory (NEI) Input Format (NIF).

Building on this data modeling, BAQ embraced the use of automated tools through participation in the EIIP Electronic Data Interchange (EDI) system prototype project. The Bureau recognized the potential of electronic data exchange to reduce the reporting burden on both industry and AQ staff while increasing the quality of data. Based on their involvement in the EIIP project, BAQ developed its own customized EDI data standard and the technical infrastructure to support reporting of the DEP Annual Emission Statement (AES).

Historically facilities reported emissions inventory data on a series of paper forms, signed and dated the forms, and mailed the forms to BAQ staff for processing. BAQ then performed varying, non-standardized checks of the data and entered it manually into the Bureau's AIMS database – a laborious and time consuming task. This process involved redundant data entry that increased the potential for human error and slowed down the overall reporting process. The Bureau knew that this process could be streamlined through automation.

Over time, as the trends in information technology (IT) tools and techniques evolved, both the EPA and DEP changed their focus for automated data transfer from EDI to the use of XML. The Bureau transitioned from their EDI prototype to the creation of a prototype XML solution. Although the initial implementation did not attract many users, it proved to be a viable means of collecting data from facilities and it provided the foundation for further expansion of the Bureau's web-based capabilities.

More importantly, through their XML prototype and the lessons learned through their involvement in the EDI prototypes, BAQ recognized that a more simplified, user friendly, accessible solution was needed. For some time, staff at BAQ had discussed the possibility of an actual online application that could be accessed locally by Commonwealth facilities, provided custom access to facility-specific data, was intuitive to use, and easy to maintain. In 2002 BAQ once again refocused its efforts and resources toward developing solutions that leveraged the flexibility and accessibility of the web, while maintaining data quality and integrity. The Bureau embarked on an aggressive project to define, design, build, and implement a full web application <u>and</u> a full XML application for the collection and review of air emission data. At the same time, they began efforts to enhance their capabilities for reporting the resultant data to the EPA under the National Emission Inventory (NEI) and Consolidated Emission Reporting Rule (CERR).

The Bureau's historic project approach relies on strong planning, open communication, and structured management processes. In keeping with this approach, their first steps toward the development of the *AES* web applications was the requirement for the use of a full system lifecycle methodology and execution of team-driven project management. Following the planned approach, the initial work was characterized primarily by completing the requirements and design phases for the online web application. The requirements and design that would become *AES*Online* (and eventually *AES*XML)* were functionally straightforward but technically complex. The requirements for *AES*Online* and the upcoming *AES*XML* application specifically met Pennsylvania's air regulations while still capturing the fundamental data needed for reporting to the EPA and other air quality management organizations.

To fully support the reporting requirements, the basic premise for the application's functional design was two fold:
1. Design the web pages to be intuitive and flexible and to resemble the existing hard copy emission reporting forms thereby providing users with continuity and familiarity.
2. Facilitate navigation and ease of use within the application by providing multiple ways of moving from page to page without being locked into a particular order of completion.

Underlying both of these requirements was the need to perform immediate and automated data validations, produce a final, human readable report for facility certification, and provide reliable performance.

Based on the completion of the application requirements definition and design, development of *AES*Online* was completed for its pilot implementation to support the 2004 AES reporting cycle. In the six months between July and December 2003, the CIBER, Inc. Team of four developers, technical writer, and project manager delivered the fully functional *AES*Online* application. This highly featured web application met the original design by better than 95 %. Additionally, the pilot participation greatly exceeded the Bureau's expectations. The initial expectation for pilot implementation was for approximately 30 facilities across the Commonwealth to successfully enter and submit AES data through *AES*Online*. At the completion of the 2004 reporting cycle, almost 300 facilities officially submitted their data. From the perspective of time saving, the reporting process time span for an individual facility Report was generally reduced from three weeks to three days. This reduction in turnaround time is further heightened by a reduction in the amount of error correction that was necessary upon loading the data into the production database. From these points of reference, the *AES*Online* application exceeded expectations in all areas.

Following the successful pilot implementation, *AES*Online* was opened for full, state-wide use in January 2005. Several enhancements were made to the application based on its first year of use, but the functional aspects of the application remained largely unchanged.

After the rapid and successful creation of *AES*Online*, BAQ commissioned the CIBER Team to design and develop a companion solution for reporting emission data in XML format. *AES*XML* development began in July 2004 and was completed in December for pilot use in January 2005. In preparation for building the XML-based solution, the AES XML Data Type Definition (DTD) created for the earlier prototype was also reengineered to create a more functional AES XML Schema.

With *AES*Online* and *AES*XML*, BAQ achieved all of its primary goals. These applications meet high standards for performance and flexibility, while featuring functionality that is robust, comprehensive, and mature. These two applications are actually more accurately characterized as a suite of three individual applications that work in concert to provide two systems for data collection coupled with a common application for data review and acceptance. This integrated approach creates two significant positive results. First, the applications share the same 'look and feel'. This makes it much easier to understand each application, eases training, and makes them generally more usable. Second, from a technical perspective, the similarities allow for the same system objects (e.g., code, database structures, infrastructure, etc.) to be shared. This reduces the cost of repeated development and makes each application easier to maintain over time.

The *AES*Online* and *AES*XML* applications are based on DEP's Oracle 9i database management system. The architecture relies on an intermediate database for direct application interface. Behind the intermediate database is the production AIMS database which is BAQ's main repository for inventory data. The graphical user interfaces (web pages) of the applications were built using Java and Oracle's JDeveloper tools. The applications are individually deployed on an Oracle 10g Application Server. Execution of the data validation and management logic is primarily built with Java, PL/SQL, XML, XSL, and HTML. To support the robust data validation and navigational flexibility inherent in the systems, this logic performs at both the page level and as a middle tier between the pages and databases.

Features and Functions Shared Across All the Online *AES* Applications
Users of *AES*Online* and *AES*XML* can complete the facility-specific AES Report online using any PC that is connected to the Internet and submit it securely to BAQ. Both the *AES*Online* and *AES*XML* applications, including the *AQ Application* which is the third, integrated application that is used by BAQ to review the submitted data online, share a common access point, similar user security structures, similar page layout and navigation motifs, and several common application features, including:
- Registration process (not necessary for the *AQ Application* as they are State employees)
- Secure login and application access
- Trading Partner Agreement policy (also not necessary for the *AQ Application*)
- A 'NOTE' system to communicate information to and from BAQ staff
- Ability to 'bookmark' the users location within the application
- Electronic document attachment capability
- Integrated online User Help
- Automated data validation and data checking
- Universal page layout, navigation, and links
- Creation of an Adobe Portable Document Format (PDF or .pdf file) of the AES Report to comply with the Commonwealth's online submission guidelines

<u>User Registration and Application Access</u>
By law, facilities must report annual emissions data on or before March 1 of the next calendar year. For electronic reporting using the *AES*Online* or *AES*XML* applications, this deadline is considered to be

met when a facility successfully completes the initial submission process on or before March 1. The timing associated with the post-submission review process will vary depending on the quality of the submission and the general workload associated with the review cycle.

Data submission through *AES*Online* and *AES*XML* is completely secure. Access to the applications is restricted by user name and password. To obtain the necessary access credentials, each facility user must complete the registration process and be approved as a user by DEP. DEP provides each user with a unique username and password. These security features not only give users access to the system, but because each user is tied directly to their facility, the security triggers the customized, facility-specific pages needed for accurate data entry and electronic data submission.

Access to the *AES*Online* and *AES*XML* applications is gained though a single point – DEP's web portal, the *DEP GreenPort*. This web site is the launching pad for each of the AES systems. Based on a successful login, users are able to navigate to the BAQ AES applications for which they are registered.

A key feature of each application is the required Trading Partner Agreement (TPA). Each user wishing to use one of the Bureau's systems for electronic submission of their inventory data must accept the TPA. Based on Commonwealth electronic reporting regulations, if a user does not agree to accept the DEP TPA, they cannot use the system to submit the AES Report.

To help users and BAQ staff with following the submission process, as an online or XML AES Report moves through the data entry/upload and review processes, the system will automatically assign a status to the Report based on its stage of processing. In addition, the Report status also determines who may perform certain operations on the Report. There are six status that a report may be identified with over the course of the submission process, including: Open (where each report starts), In-Progress (while data entry or .xml file preparation is underway), Completed (data is pending submission by the facility responsible official), Submitted (officially transmitted to DEP and pending review), Revision (errors have been identified by BAQ and are being addressed by the facility), and Accepted (where each report ends).

Creation of Notes

Both *AES*Online* and AES*XML* provide a convenient method to add notes to a facility's automated AES Report that might explain data, request assistance, or otherwise communicate details about the report to BAQ. The *AQ Application* also provides BAQ staff with a means of creating 'memoranda to file'. The Internal Memo page provides a method of communication between AQ staff and also provides a method to maintain internal information pertinent to a given report. The internal memo feature is only available to AQ users. Facility users cannot enter an internal memo and do not have access to the notes entered by AQ Reviewers on the Internal Memo page.

Bookmarks

Throughout the *AES*Online* application and the *AQ Application* users may insert a bookmark on any page in the application. Users are able to bookmark as many data entry pages as they wish, thereby supporting the flexible data entry flow and for reentry to the systems.

Document Attachments

In addition to the standard AES Report sections, *AES*Online* and *AES*XML* also provide the ability to "attach" electronic versions of supporting documentation. The systems limit the size of an individual attachment to 5 MB and the combined size of all attachments for a given Report to 30 MB in order to conserve file storage space and bandwidth. Also, to prevent the uploading of malicious files and limit the proliferation of viruses files being attached are limited to:

- Word Documents (.doc)
- PowerPoint Presentations (.ppt)
- Excel Spreadsheets (.xls)
- Text or Rich Text Documents (txt or rtf)
- Common Image formats (.bmp, .gif, .jpg, etc.)

Examples of this supporting documentation include documents or spreadsheets containing emission calculations or forms specific to a DEP region.

Online User Help

Each of the *AES* applications includes an integrated, online User Help system complete with Search capability. This Help system provides page and field-level information to assist users in completing, revising, and submitting the AES Report. Each page in the applications provides a direct link to online Help for that page and a method to contact appropriate BAQ staff for additional assistance. In addition to online Help, the applications also provide off-line user help that can be downloaded to the user's PC or printed.

Data Validation

Perhaps the most powerful part of the applications is the data validation that occurs. When a facility user attempts to submit a report through *AES*Online* or upload an .xml file through *AES*XML*, the systems automatically perform numerous validation checks on the data before it is transferred. Data are validated against established criteria to ensure that the file is complete and the data are valid and within acceptable ranges.

For *AES*Online*, these validations are performed at two points. As data is being entered, specific automated validation functions are performed on each page throughout the application to ensure completeness of the final dataset. Completeness at the data element level is verified by simple checks of whether required data elements have been entered. Beyond checking at that level, some of essential data being entered by the facility is compared with data maintained by BAQ based on their individual permit and previous year's data submission. This 'static' data such as the sub facilities (operating units) list, SCCs, and contact information, is preloaded into the system at the start of the reporting cycle. By doing this, the *AES* applications are able to provide custom, facility-specific pages based on the user's login and can perform this heightened level of data comparison and automated quality assurance.

If conditions are violated at the system-level validations, users must correct the error, typically missing data, before being allowed to move to the next page or exit. Once all the required data has been entered and the report is considered by the facility to be complete, the user may initiate the submission process. When a facility user initiates the *AES*Online* submission process, the application checks all components of the pending AES Report to ensure that the required emissions data has been provided. The *Submission Checklist* page displays a checklist of these required components, indicates whether they are complete or incomplete, and provides links to the incomplete items. All items on the checklist must be completed for the submission process to continue. This second level of validation includes dozens of range, unit of measure, and regulatory threshold checks against data related to:

- Fuel Usage Summaries
- Facility (Site)–Level Manual Emissions
- FML Fuel Tests
- SCC "Local" Fuel Tests
- Operating Schedules & Throughputs
- Sub Facility Manual Emissions
- Miscellaneous Emissions
- Various other completeness criteria

By reducing data errors in the submitted Report, these system checks improve the chances that the initial submission of the AES Report will be accepted. In the past, hardcopy submission required significant

effort to ensure completeness and accuracy. Based on feedback from the user community, the paper-driven process that once took weeks of preparation and transcription has been reduced to days and, in some cases, hours.

Within *AES*XML*, once a file is uploaded and the data submission process has been initiated a series of automatic data checks are launched to ensure that the .xml file is complete and that the data is valid and within acceptable ranges. The data set is verified for its formation against the AES XML Schema (Schema Validation) and against the BAQ business rules (Extended Validation). The Schema Validation check verifies that the .xml file being submitted is compliant with basic rules of XML structure. If any data item fails this check, the user is provided with a page indicating the upload failure and the reason for the failure. The issues identified must be remedied and a corrected file uploaded in order for the upload process to continue.

Once the schema validation is passed, the Extended Checks are triggered. *AES*XML* validates the data within the .xml file against criteria that are generally the same as those included in the *Submission Checklist* and at the page-level within *AES*Online*. Again, if any data item fails the checks, the system displays a page characterizing the errors. The facility user must resolve all of the identified issues and upload a revised file prior to continuing the submission process. An .xml file cannot be uploaded into *AES*XML* unless it passes all of the data validation tests.

As with *AES*Online*, the implementation of these robust data checks creates data sets that are readily usable by the Bureau. Depending on preparedness of a facility using *AES*XML*, once a facility user is set up to generate an AES .xml file, the robust validation processes and automatic data transfer further reduce the overall processing time from the weeks witnessed with the paper process and days associated with *AES*Online*, to a matter of hours or even minutes.

Page Layout, Navigation, and Links
Beyond the commonalities found between *AES*Online* and *AES*XML* relative to the overall process, each application shares common page layout, navigation, and information link attributes. Each page in the applications includes a menu bar with links to other pages in the system. This menu bar appears in the blue area on the left side of each page.

The Left Menu Bar provides quick and simple access to all of the data and data entry pages needed to complete a facility's AES Report. The layout and organization of this menu is similar to that of the paper form of the AES. Each menu selection is a link to a different application page. The navigation menu is designed to allow users to navigate to any section of the online AES Report in a few clicks as possible. In most cases a particular page can be accessed by only one or two mouse clicks.

In addition to the Left Menu Bar, in the footer section of most pages in the applications, links are provided to the DEP Home Page, DEP contact information, online Help, DEP help desk, and to exit the system.

Non-Automated Activities
Although *AES*Online* is a comprehensive automated system, certain procedures remain outside the scope of the *AES*Online* system and must be handled manually by DEP and BAQ staff. Specifically, facility users must contact BAQ to:
- Edit or view previous year's data.
- Create a new primary facility or sub facility.
- Modify existing sub facility maps and SCCs.
- Modify a sub facility confidential data indicator (apply confidential status to data or remove confidential status from data)
- Modify User Access or User Security Roles.

- Change AES Report submission preference from *AES*Online* or *AES*XML* to paper submission forms.
- Modify data in an AES Report that was submitted and accepted by BAQ and is in "Accepted" status.

Overview of the Online *AES* Applications

There is core information in common between *AES*Online*, *AES*XML*, and the *AQ Application*. Although the actual pages that are presented within the individual applications may vary slightly, users of either application are presented with data and information that have been identified as being of significant importance for aiding in preparation of the report and for ensuring a quality submission.

- ***DEP Contacts Page*** - displays DEP contact information related to the facility submitting the AES Report. The contacts are tailored to the individual facility based on permit information and historically submitted information.
- ***Facility List Page*** - the first page in the *AES*Online* and *AES*XML* applications. This page lists all the facilities an individual user is authorized to access.
- ***Facility Details Page*** - displays information about the primary facility's identification numbers, location, and contact information. The displayed information is populated automatically from previous AES Reports submitted by the facility. If the facility needs to modify any of this information, they are advised to contact BAQ immediately to request the necessary changes.
- ***Sub Facility List Pages*** - lists all of the primary facility's sub facilities (operating units), sorted by type (Fuel Material Location, Combustion Unit, Process, Incinerator, and Control Device).
- ***Sub Facility and SCC Status Change Request Page*** - used to request a change in the status of an SCC. When an SCC status change request is submitted, the application generates e-mails confirming the request and sends the request automatically to the corresponding BAQ and facility users. Once received the responsible BAQ staff review the request with the facility and make the appropriate adjustments. This feature streamlines the communication process, but does not allow the facility to unilaterally make changes to its regulated information.
- ***Sub Facility Maps Page*** - provides users with a graphical representation of the connections between sub facilities, including all combustion units, fuel material locations, control devices, and emission points included in the process flow. For example, a specific FML provides fuel to the following combustion units, which in turn feed this list of control devices, and the emissions go to these specific stacks.
- ***Emission Factors Page*** - displays EPA emission factor information, listed by SCC. This information page is provided to assist facility users with determining emission amounts for their specific sub facilities.
- ***View PDF Page*** - as part of the submission process at the facility, the *AES*Online/AES*XML* system generates a .pdf (PDF) file of the facility's Report. This file provides a "snapshot" of the Report data in a readable and printable format that cannot be edited. Each time the Report is submitted, a new PDF file is generated, recorded in the system, and added to the PDF list on the *View PDF* page. Only the most recently submitted PDF file will be available in the system for viewing for any year. Within the *AQ Application*, BAQ staff are provided with the last PDF file submitted by the facility for the current reporting year and the PDF files on record for the facility for the previous two reporting years, if available.

AES*Online – The Application for Facility Online Data Entry

The previous sections described the features and functionality that are held in common between the *AES*Online*, *AES*XML*, and the *AQ Application* that is integrated with both of them. The scope of the common features, including the rigorous data validation logic, online user help, and data review pages comprise a significant amount of functionality in themselves – enough to make a relatively complete application. This section highlights the application features and functions that are unique to *AES*Online*.

Take together, the common features described above and the unique system components being described here come together to form *AES*Online*.

As *AES*Online* is a true online application, it allows users to record and submit facility emission inventory data through a series of customized web pages presented on the Internet. The main user group that might use this system are those that do not have a single data source for their AES data or who do not manage their data in a more automated manner. These are the users that perennially transcribe their data from various sources to the hardcopy forms and do not otherwise manage their data electronically. For ease of transition from paper to web forms, the *AES*Online* pages are designed with names and layouts similar to the paper version of the current AES Report. Users are able to completely enter their AES data by navigating the system from a navigation frame that is present on the left side of each page and is customized for each facility. The system provides separate pages for entering fuel usage, fuel tests, operating schedules and throughputs, manual emissions, miscellaneous emissions, and pollutant emission totals for the facility.

Beyond the common pages and features that are shared across the *AES* applications, *AES*Online* reflects customized, facility-specific data entry pages, pages for reviewing facility and sub facility details maintained by DEP based on permits and previous AES data submissions, and customized, error-specific revision pages. To support these specialized features the following pages are unique to *AES*Online*:

- **Home Pages** - The *Open/In Progress Home Page* provides a starting point to access all the pages a facility needs to enter or modify a primary facility's AES data, and submit the AES Report to BAQ. As the online AES submission moves through the various status changes associated with the submission process, a number of other *Home Pages* are provided. These pages display different information and provide different functionality depending upon the submission activities that may be performed.
- **Pollutant Summary Page** - displays a list of all pollutants emitted by the primary facility and the annual emission amounts for each pollutant. Facility users use this page to enter and edit data. They may also add pollutant(s) to the list, enter and edit pollutant data, or remove pollutant(s) from the list. The pollutants reported in the primary facility's previous AES Report, as well as any new pollutants entered on the *Manual Emissions* or *Miscellaneous Emissions* pages for the facility's current AES Report, are included automatically on this page.
- **Fuel Usage Summary Page** - used to enter and edit a primary facility's fuel usage data. The system lists the types of fuel reported on the facility's previous AES Report, including both FMLs and SCCs (local fuels). The facility user may also add a fuel type and associated usage amounts; edit usage amounts or units of measure; or remove a fuel type that is no longer used by the facility from this page. The system will automatically validate the fuel usage amounts to ensure that no value is more than 50% greater than the sum of all throughputs for the fuel.
- **Fuel Usage Summary by SCC Page** - displays the facility's total fuel usage amount by SCC, further broken down by the totals for each sub facility**.**
- **SCC Details Pages** - The application provides a set of sub facility-specific pages that show the details of the related SCCs. There are pages for local fuel sources (a fuel source used only by that sub facility) and fuel sources associated with fuel material locations (a stand alone fuel source that may be used by multiple sub facilities). For local fuel sources, associated fuel and fuel test data can be entered or edited from this page. The FML fuel source page displays data related to operating schedule and fuel usage.
- **Add New Operating Schedule and Throughputs Pages** - is used to add a new operating schedule for a sub facility, including data on how long the sub facility operated during a given time frame and the corresponding fuel throughputs.
- **Operating Schedule and Throughputs Pages** - displays operating schedule data for a sub facility, including data on how long the sub facility operated during a given time frame and the corresponding fuel throughputs. User can edit or delete an operating schedule and associated

throughput data, but cannot add a new schedule on this page. To add a new schedule, they must use the provided *Add New Operating Schedules and Throughputs*.

- **Fuel Material Characteristics/Fuel Tests Pages** - record fuel characteristics and fuel test data for a Fuel Material Location (FML). Data can be entered or edited on this page. Since fuel material characteristics and fuel test data for an FML apply to all sub facilities that are fed by the FML, the data does not have to be entered separately for each sub facility. This page also records multiple fuel test data for each FML.
- **Manual Emissions Page** - used to enter or edit data for pollutant manual (actual) emission estimates for a sub facility. Pollutants and manual emission data can be added to or removed from the list of pollutants for the sub facility. Users also provide BAQ with the calculation method used to derive the emission amounts entered, including, Continuous Emission Monitoring (CEM) Devices, Stack Tests, or calculation-based EPA factors or control device efficiencies. The user may also choose to have BAQ use the EPA factor for calculating their emission values.
- **Miscellaneous Emissions Page** - a sub facility may emit significant amounts of pollutants that are not listed on the facility's permit and therefore not on record with DEP. *AES*Online* provides a *Miscellaneous Emissions* page to record these pollutant emission amounts and also to capture any facility-level emission amounts that cannot be attributed to a single sub facility.

AES*XML – The Application for Direct Data Submission

As with *AES*Online*, *AES*XML* is characterized by well developed, full featured, and unique functionality. *AES*XML* is the newest component of DEP's suite of emissions inventory applications. It provides Pennsylvania facilities with a fully-automated, web-based system to submit AES data in .xml file format directly from their host data source to DEP. *AES*XML* is most appropriate for Pennsylvania facilities that already use automated or semi-automated methods of AES data collection and reporting.

*AES*XML* allows for the direct transfer of data from the facility's data sources to the DEP databases. The necessity of data entry is eliminated. Although a series of customized web pages are provided to support the creation of the .xml files, the process for submission is simply focused on uploading a file through a single web page. As highlighted above in the section describing the common features of the solutions, there are many similarities between the pages of *AES*XML* and *AES*Online*. This section will outline the features and functions that are unique to *AES*XML*. Together, the common and unique features form the overall AES*XML application.

Preparation by the Facility

As *AES*XML* is driven by the submission of an .xml file, a facility using this application must have two basic capabilities; 1) a data source for the AES data, and 2) a means of creating an .xml file. In general users of this application will maintain their AES data in a database system or even in spreadsheets. Based on the database or spreadsheet system they maintain, facility users must create a valid .xml file by mapping their raw data to the AES XML schema provided by the Bureau. Beyond the raw emission inventory data that the facility would routinely maintain, BAQ also provides each registered facility with the required code lists, schema, and data dictionary needed to report through the system.

When a facility user has a structured data source and the capability to generate a .xml file, the *AES*XML* application give the user the ability to upload the file, add notes, attach any supporting documents, and submit its *AES*XML* Report online in just a few simple steps.

Beyond the common pages and features that are shared across the *AES* applications, *AES*XML* comprises pages for uploading AES .xml files, pages for reviewing facility and sub facility details maintained by DEP based on permits and previous AES data submissions, facility-specific pages for downloading DEP code lists, and customized, error notes pages. To support these specialized features the following pages are unique to *AES*XML*:

- *Home Pages* - - The *Open/In Progress Home Page* provides a starting point to access all the pages a facility user needs to enter notes, attach documents, upload an .xml file, and submit the *AES*XML* Report to BAQ. This page is the launching point for uploading an individual facility's data and contains links to data and information that is absolutely critical for a successful XML submission. The page contains links to the Bureau's AES XML schema and associated data dictionary. This is a single reliable source for the current AES XML schema in use by BAQ. The facility user may download the schema and associated descriptive documentation from this page. This page also has a link to an example .xml file that may be useful in creating a valid file. Additionally, this page has links to the various code lists are custom created for the facility. These codes must be used within their data set in order to pass data validation steps. As with *AES*Online*, as the online AES submission moves through the various status changes associated with the submission process, a number of other *Home Pages* are provided. These pages display different information and provide different functionality depending upon the submission activities that may be performed.
- *Download Code Tables Page* - provides a facility-specific code tables in CSV (Comma Separated Values) or XML (eXtensible Markup Language) file formats for download. These code lists include calculation method codes, units of measure, pollutants, fuel types, sub facility list (facility-specific), and SCCs by sub facility (facility-specific)
- *File Submission Pages* - *AES*XML* provides the user with a set of pages for the upload of their .xml file, including pages reflecting validation errors, upload failure, and upload success. The submission process is initiated by uploading a valid .xml file.
  - *Upload XML File Page* - The *Upload XML File* page allows the facility user to browse their own file system for the .xml file that they generated for submission. This is a normal, windows-based browse feature and the user may select their file and initiate its uploading to *AES*XML* from this page.
  - *XML File Upload Failed Page* - displays if an attempted .xml file upload fails one or more of the schema-level validations that are automatically performed. This page provides a list of the errors that must be addressed before the file can be successfully uploaded. The facility user must address the issues through changes to their data within their originating data source or in their .xml file generation tool.
  - *XML File Error Messages Pages* - displays if an attempted .xml file upload fails one or more extended validation checks that are automatically performed if the file passes the schema validations. This page provides a list of the errors that must be addressed before the file can be successfully uploaded. In most cases, errors reported through these checks must be addressed by the facility user through adjustments to their data within their originating data source.
- *Report Rejection and Resubmission Pages* – If BAQ rejects an AES Report submitted through *AES*XML* the facility user receives an email indicating the rejection and identified issues. At that time, the status of the report is set to Revision and the facility users must address each of the issues and resubmit the report.
  - *Error Notes Menu Page* - upon receipt of the 'rejection' email, if the facility logs onto *AES*XML* again, in addition to all of the pages that were available when the Report was in In-Progress status, they will also have access to the *Error Notes* page. This page provides additional information from BAQ about the errors or questionable data that were found through the BAQ analysis. In most cases, revision of the data will require modification to the facility's data within its originating data source. Upon verification or revision of the data, the facility user must generate a complete, new .xml dataset for resubmission. The pages that they use to upload the revised .xml file and resubmit the AES Report perform the same functions (validations) as the pages used to complete and submit the initial Report.

AES*Online/AES*XML AQ Application – The Application for AQ Data Review

The *AES*Online* and *AES*XML* applications actually share a third, integrated system – the *AES*Online/AES*XML AQ Application*. The *AES*Online* and *AES*XML* applications are external systems which the industries use. The *AQ* or *Internal Application* is an internal system used by BAQ staff for the review and acceptance of the submitted data. Although it is an 'internal' application, it is still an Internet application that is accessible from anywhere there is Internet access. As such, the common page features and access functions described above for *AES*Online* and *AES*XML* are generally shared.

AQ User Security Roles
The allocation of access credentials and login requirements for the *AQ Application* is slightly different than the other *AES* applications. The assignment of user names and passwords is not necessary because only DEP staff access this application and they have pre-existing security credentials.

The security roles for the *AQ Application* are slightly different from those for *AES*Online* or *AES*XML* as well. All AQ staff that have access to *AQ Application* are assigned the user security role of Reviewer. Assigned Reviewers can request verification of data, create enter notes to the facility, create internal notes to the file, create error notes and Accept or Reject the report. They may not edit any of the data. For Reports in Open, In Progress, Completed, Revision, or Accepted Reports, Reviewers can only view the Report home page, PDF file, and any associated notes and attached documents.

Problem Sub Facilities Page
The *Problem Sub Facilities* page provides BAQ Reviewers with a list of an individual facility's sub facilities that include at least one error icon. This page helps Reviewers in quickly identifying sub facilities with potential issues and navigating immediately to the source of the listed error icon(s), thereby accelerating the data analysis process.

Data Comparison Reports
At the heart of the *AQ Application*, beyond the automatically generated error icons, are the various comparison reports that are available. When a facility user submits the AES Report, the application automatically generates a suite of reports that compare the submitted data against the same data reported for previous years. The intent of these reports is to highlight any anomalies in the data that might be occurring from year to year, thereby increasing data quality and accelerating the overall data review process.

- ***Facility Totals Comparison Report*** - provides a comparison of the facility's total manual emissions data submitted for the current year with the total manual emissions on record for the previous reporting year. The facility's manual emissions data reported two and three years previously are also provided. Error notes may be generated by the Reviewer from this report and result in data flagged for review by the facility upon rejection of the report. The report lists all pollutants used by the primary facility with the criteria pollutants being highlighted. Pollutant manual emission amounts that differ from the previous year's amounts by more than 10% are flagged as being questionable.
- ***Sub Facility Totals Comparison Report*** - provides a comparison of the facility's total sub facility manual emissions data submitted for the current year with the total sub facility manual emissions on record for the previous year. Emissions data reported two years previously are also provided. The report lists all pollutants used by the sub facilities with the criteria pollutants being highlighted Pollutant manual emission amounts that differ from the previous year's amounts by more than 10% will be flagged as being questionable.
- ***Sub Facility Totals Comparison Detail Report by Pollutant*** - provides a comparison of current and previous year's manual emission totals for a selected pollutant, by sub facility. Error notes can be assigned, edited, or cleared on this page.
- ***Operating Schedule Comparison Report*** - provides a comparison of the total annual hours of operation reported for the current reporting year with the total annual hours of operation reported

for the previous two reporting year by sub facility and SCC. The *Operating Schedule Comparison* page can be displayed for a specific sub facility and the display can also be order by the percentage difference from the previous year.

- ***Throughput Comparison Report*** - provides a comparison of the total annual throughput amounts reported for the current reporting year with the total annual throughput amounts reported for the previous two reporting year by sub facility and SCC. The *Throughput Comparison* page can be displayed for a specific sub facility and the display can also order by the percentage difference from the previous year.

Overview of the *AES* Submission Process

After online data entry into the *AES*Online* application or a .xml file upload into *AES*XML*, the submission process involves clicking a few buttons and following simple instructions. By using the online systems, users no longer have to e-mail or mail paper forms - the system transmits the data with the click of a mouse. During the submission process, the system generates an Adobe Portable Document Format (PDF) version of the specific facility's online AES Report. The PDF file is a snapshot of the AES Report data that was entered for the facility. The PDF becomes part of the official DEP record. The PDF version of the Report is displayed in a format similar to the paper forms with which the facility users are already familiar. If the facility notices any errors or discrepancies in the PDF report through own manual review, they may return to the online pages and continue entering and editing data. If the PDF is accurate, the report may be submitted to BAQ for review. By reducing data errors in the submitted Report, these system checks improve the chances that the submitted AES Report will be accepted.

Each facility Submitter, or Responsible Official, must review the PDF and certify that it is an accurate representation of the emissions inventory data before the system will allow the AES Report to be officially submitted to BAQ for review. This step is necessary to satisfy the Commonwealth's requirements for electronic submissions. It informs the user of the provisions of the PA Electronic Submissions Act (Act 69) and confirms the user's intention to submit the AES Report to DEP.

After a user successfully submits an AES Report to DEP through either application, e-mails confirming the submission are automatically distributed to all registered users at the facility, to the BAQ Reviewer, and to the designated BAQ supervisor. This communication confirms the completion of the data transfer and alerts BAQ that data is available for review. Users are also presented with a page that confirms the submission of the AES data. More importantly, the email message provides the Submission ID which is essential for communication between the facility user and BAQ staff.

When the AES Report is submitted online the data is immediately available to BAQ for review, analysis, and acceptance or rejection. To help aid the analysis and review of the data, immediately upon submission the applications perform further data comparisons. These comparisons automatically identify (flag) any potential inconsistencies, discrepancies, or problems based on comparison with previous years data.

The BAQ Reviewer assigned to the individual facility accesses the submitted Report through the *AQ Application*, reviews any system-generated error flags, analyzes the system generate comparison reports, and performs additional reviews of the data for accuracy and completeness based on their knowledge of the facility's annual operation. The BAQ Reviewer will accept or reject the submitted Report depending on the results of this review. If errors are identified, the Reviewer will use error icons to flag the data for the facility's attention and reject the Report pending revision.

To help BAQ with communicating errors within a facility's AES Report, the *AQ Application* uses a system of icons to flag questionable or invalid data and to alert facility users to potential errors or

modified data. Some of these icons are automatically generated by the system, some are assigned by the Reviewer, and some are generated by the facility user.

When a BAQ Reviewer analyzes a report that has been submitted for the first time, non-critical error icons that are generated automatically by the system may be displayed. These icons are considered 'non-critical' and will not prevent a Report from being accepted. They provide reference information and may alert the Reviewer to potentially erroneous data, but do not by themselves indicate an actual error.

- *Questionable* icon - is automatically assigned by the system to any data that fails BAQ-defined tolerance limits set for the data. Questionable data may be correct, and are evaluated by the BAQ Reviewer for potential errors. The presence of one or more *Questionable* icons in a report does not prevent its acceptance. The Reviewer may accept a report even if it includes data items flagged with the *Questionable* icon.
- *No Data to Compare* icon - is automatically assigned by the system to emissions data if the current year's reported data is compared with the previous year's reported data, and either no record of the data element exists for the previous year or the recorded data value was "NULL."
- *AQ Action* icon - is automatically assigned to each data element in the Report that is editable by the facility and not flagged with a *Questionable* or *No Data to Compare* error icon. The *AQ Action* icon does not indicate an error or questionable item, but provides a method for the Reviewer to request verification (assign a *Verify* icon) to data that was not flagged by the system as potentially invalid. This icon is considered a "non-critical" icon and has no effect on the acceptance or rejection process if it left unchanged. If a Reviewer wishes verification to a data element that has this icon, they simply change it to a *Verify* icon and enter an error note.

The system also uses a number of 'critical error icons' to alert users to data that may be erroneous or may have been modified. These icons may result from an action on the part of a BAQ or facility user. A report that contains one or more critical error icons cannot be accepted. All critical error icons must be cleared before the system will accept the report.

As a BAQ Review analyzes a submitted report, they may identify data that needs to be further addressed by the facility. In these cases, the BAQ user can select one of the non-critical error icons that are displayed, enter a note about the potential error or question, and request verification of the data. The system will automatically change the original icon (one of the non-critical error icons) to the *Verify* icon. This will generate a 'critical error icon' for the facility, the Report will be rejected, and the issue will be addressed by the facility user.

If BAQ rejects a report submitted through *AES*Online* or *AES*XML*, the facility user will receive an e-mail confirming the rejection and noting the reason(s) the report was not accepted. The status of the submitted AES Report will be changed to Revision. For users of *AES*Online*, edits are made directly through the system in the same manner used for the initial data entry. For *AES*XML*, users must make any necessary adjustment in their data source, regenerate a new, complete .xml file, upload it to the system, and resubmit the AES Report to DEP. For both applications, this process can be repeated as many times as is necessary to achieve a complete, error-free data submission that can be accepted by DEP.

To help the revision process, three error icons may be displayed to the facility users on their Reports in Revision status.

- *Verify* icon - is assigned by the BAQ Reviewer during the review process to any questionable or incorrect data. When a rejected report is received, the presence of this icon will alert the facility user to correct the error or verify the data. This is the most important error icon in the system as it indicates a data item that must be specifically addressed and can not be ignored. When a facility user accesses the report in Revision status the only icons that they will

see are the *Verify* icons. If they move the mouse over a *Verify* icon, any error note that was entered by the Reviewer for that error will be displayed.

- *Acknowledged* ⚠ icon - indicates that the facility user has verified or updated information that was initially flagged by BAQ with the *Verify* icon. When data is changed or verified by the facility user, the system automatically changes the icon from *Verify* to *Acknowledged.* This icon will also be displayed to the BAQ Reviewer upon resubmission. *Acknowledged* icons are considered "critical" error icons. A facility user cannot resubmit a report until all the data flagged with *Verify* icons have been addressed (changed to *Acknowledged* icons). Similarly, all *Acknowledged* icons must be cleared from a Report by the BAQ Reviewer before it can be accepted or rejected.
- *Modified* 📝 icon - identifies a data item that was not originally flagged with a *Verify* icon by BAQ staff but was modified by the facility user while the Report was in Revision status (during the last set of revisions). This icon will also be displayed to the BAQ Reviewer upon resubmission and helps BAQ staff identify any differences between the revised and original submissions.

Once a facility's AES Report has been accepted by BAQ, the report status is changed to Accepted and it becomes the facility's official submission to meet State and Federal regulations. At that time an email message is automatically generated from the system and sent to all individuals registered at the facility. No further edits or submissions are necessary through the *AES*Online* or *AES*XML* applications from that point. Facility users may access the system at anytime to view their submitted PDF, notes, or other static information, but they may not perform any data entry.

Once the Report is accepted, the data is automatically imported directly into the Bureau's enterprise-level database – AIMS. Since this data is already in the system, DEP staff do not need to reenter any Report data. This reduces the DEP workload and also greatly reduces the opportunity for data entry errors. Any issues that may be identified with the submitted data at this point are resolved directly with the facility outside of any of the online systems.

Using *AES*Online* and *AES*XML*, the entire AES reporting and review process is accomplished over the Internet. Although *AES*Online* and *AES*XML* are automated and more efficient than the paper form process that has been used historically, the same level of personal communication shared between industry and BAQ is still enjoyed. Links to contact information for BAQ staff, including e-mail addresses and telephone numbers, are readily available on each application page. If facilities have questions, they can still call and speak person-to-person with a knowledgeable DEP staff member.

CONCLUSION
Although this paper has offered a detailed description of the web-based systems that the Pennsylvania Bureau of Air Quality's has built for the collection of air emission inventory data, the main point is that the Bureau established many significant and difficult goals and accomplished them all. The Bureau persevered through evolving requirements, changing IT approaches, and a couple of nearly failed prototype adventures. Rather than simply maintaining historical processes because they were known and because attempts at more sophisticated automation were less than sterling, they continued to focus on an end result that would:

- Streamline the emission inventory process
- Increase data quality, access, and use
- Save costly preparation and processing time
- Cut recurring financial and resource costs
- Ease reporting burden
- Refocus staff on functions beyond data entry
- Reduce paper use, postage, and filing space

In the end, the Bureau invested in and obtained a suite of robust, flexible, and intuitive applications: two applications in *AES*Online* and *AES*XML* that directly serve the reporting industries of Pennsylvania

and the *AQ Application* that greatly enhances their own ability to manage their data. What's more, these solutions are coupled with their Air Information Management System (AIMS) that is the core of their day-to-day emission inventory, permitting, and compliance data management and a suite of web applications that maintain their ability to compile, manage, and share their data with EPA under the CERR and other air quality management organizations. Taken together, BAQ has created a comprehensive system for managing their emission inventories – AND IT WORKS.

In its second year of use, *AES*Online* supported the online submission of Annual Emission Statement reports from nearly 500 facilities across the Commonwealth. That is more than a fifty percent increase over its first year of use (for the first year the Bureau desired thirty facilities to use the system and actually got nearly 300). These are very significant numbers and results given that the Bureau cannot mandate the use of electronic reporting and is limited in the amount of marketing that can be done. By mostly word of mouth, distribution of routine written communication, advocacy provided by Regional staff, and the actual results, each of the facilities used the system voluntarily.

Of the online facilities, nine were pilot participants using the AES*XML application. This group was intentionally kept small based on the more technical nature of the implementation and the historical track record of the EDI and XML prototypes. Although not a huge percentage of the overall number of reporting facilities, the nine facilities that were fully engaged in the pilot successfully completed the pilot. This is significant given that only a few had IT resources available to assist with getting their XML capabilities refined. The balance were users who were perennially responsible for compiling the AES Report and had little experience in automated data management. Each of the users displayed patience, desire, and a belief in the approach of setting up a system once so that it will bring ease for the future. Now they all have databases and spreadsheets, that are mapped to the BAQ XML schema, that successfully compiled and transferred AES data electronically, and will do so again next year in a matter of minutes without data entry beyond initial data collection (none for those with automated data collection tools).

The system has delivered. *AES*Online* and *AES*XML* are easy to access and perform well under heavy workload. Data entry and file upload are intuitive and smart enough to catch the minor errors. Feedback related to invalid data is rapid and accurate. The entire process is accelerated. From the perspective of the air quality staff the feelings are the same. For those reports submitted electronically, data entry is totally eliminated. Standardized data review is now available. Filing and paper processing is significantly reduced. The entire process is streamlined and accelerated.

In the end, the *AES*Online* and *AES*XML* applications, although they required faith and investment, have become cornerstones of the Bureau's air emission inventory management. They perform an important role in ensuring that quality data is available for use and that the level of resources required to acquire it is minimized. These solutions and their associated systems will continue to positively return investment for the Pennsylvania Department of Environmental Protection in terms of data and dollars.

KEY WORDS
Consolidated Emission Reporting Rule (CERR)
Data Type Definition (DTD)
Electronic Data Interchange (EDI)
Emission Inventory Improvement Program (EIIP)
eXtensible Markup Language (XML)
National Emission Inventory (NEI)
Trading Partner Agreement (TPA)
United States Environmental Protection Agency (USEPA)