Combined Air Emissions Webinar
Questions and Answers

1. **Could you expand on what is the "problem" being solved? I see value in the goals mentioned. That being said, how is the current process not working or addressing the reporting needs?**

   Looking at slide 8 from the webinar, our Lean event identified a number of inefficiencies. These were identified based on "root cause" analysis of some specific problems that have been noted about the air emissions report across all EPA and state/local/tribal (SLT) agencies. These problems include:

   - National Emissions Inventory (NEI) data take much longer to publish than the Toxics Release Inventory (TRI) and Greenhouse Gas (GHG) data, slowing down the various uses of this data and costing the EPA and SLTs more by dragging out the process.
     - The EPA spends a lot of time augmenting data from the states when it would be more efficient just to get the data right when it's submitted the first time.
   - Facilities must enter the same information about their facilities into multiple data systems (wasting time) and leading to inconsistencies.
   - Inconsistent emissions data across data systems reduces credibility.
   - The same quality assurance (QA) steps are being done multiple times, for example, the same "questionable" hazardous air pollutant (HAP) emissions value could be reviewed by TRI, NEI, and a SLT agency, which wastes time.
   - The emissions data quality can be improved and transparency increased by ensuring test data received by the EPA (via the Compliance Emissions Data Reporting Interface (CEDRI)) is used for computing emissions. Without including such a link, agencies are not necessarily using the "best available" data to compute emissions, but doing so is a requirement of both the NEI and TRI programs.
   - The EPA is unable to easily publish even a facility-level total of emissions across all programs because: (1) matching facilities across programs is not trivial with constantly changing IDs and other facility attributes and (2) inconsistencies in emissions across TRI and NEI make it impossible to publish a "single" value until those discrepancies are resolved.

   In addition, we have heard from our industry stakeholders that they incur costs and wasted time when inaccurate or inconsistent data about their facilities is published by the EPA. Some industry commenters have noted that for their facilities, the expense associated with defending inaccurate and/or old data is the largest problem. Thus reducing inconsistent, outdated, and inaccurate information is the most important possible improvement that this project could provide.

2. **My question about the "problem" being solved was really answered. There was no scope or scale to the "data" concerns mentioned. I'm concerned this project is "solving" a problem that isn't a problem or monies spent on other concerns.**

   Please see the answer to question 1 above. The Return on Investment (ROI) analysis will capture the scale of the problem and help to make sure that it's worth doing something about.

3. **In light of the fact that many states have well established reporting programs, what problem is being solved by this? I see that we may have to start over with our program which we have spent many years and $100,000's to build.**

   We understand that many states have made large investments and have good emissions reporting programs. Unfortunately, the problems listed in the answers to question 1 remain with continued costs for both industry and government. A holistic solution has not yet been attempted, but this is just what we are attempting now. We believe we can include the high quality state systems as a part of the "to be" solution without "starting over".

4. **With respects to the inefficiencies, can you provide a level of scale? How much duplication for example?**

   There are about 9,300 TRI facilities that are also included in the NEI for NEI pollutants. While relatively few facilities report to CEDRI at this time, the use of CEDRI will be expanding to include nearly all of the facilities (about 75,000) that report to states for use in the NEI. There are some facilities that report to all four programs, but relatively few at this time.

5. **NEI inefficiency: Are you referring just to the point source category or the whole process (clarification: Whole process meaning all categories, mobile, area, etc..)?**

   The project addresses point sources only, because that is where the industry stakeholders have the most duplication of effort. Since one of the goals of E-Enterprise is better support of the regulated community, the other NEI source categories such as mobile and area sources do not apply as much. However, the NEI program continues to identify and implement improvements to these other source categories.

6. **Our state obtains emissions data annually from facilities that have air emissions greater than 5 tons. We also have a list of air toxics recognized by the state which goes beyond the list of federal HAPs. How would EPA use a shared system to incorporate these smaller facilities and expanded pollutant list into something like the NEI which has much higher criteria air pollutant thresholds, and voluntary HAP emissions reporting?**

   The "to be" state would allow for collection of any/all air emissions data identified during any future design phase of the project, which can include even those data elements not required by the NEI or other federal reporting programs. If the same data reporting approach is used by all air agencies, including SLT air agencies, then in principle the data that goes beyond EPA's requirements would be available to a wider audience. Rather than states having to take a separate step to deliver the not-required data to the EPA, the EPA would have the same data as the state has available. Policies for how these data can be used and accessed would need to be devised, but in principle, the Combined Air Emissions "to be" state promotes easier voluntary data sharing by making it automated. Please note that the NEI already has many sources as point sources that are far below the required thresholds.

   Any facility that is regulated (or otherwise tracked) by any participating agency could be included in the Shared Facility Attributes part of the system. The question implies that perhaps just federally required sources would participate, but that is not accurate. The proposed "to be" state is not intended to replace any current state-level reporting systems, although it is possible that current systems may need to be enhanced to be able to function with the "to be" state. The advantage of

such an approach would be a mechanism by which submitted data can be more easily submitted, QA'd, and shared amongst the various reporting programs.

7. **At some point, will there be a move to eliminate/consolidate data input under a single unified ID for a facility? Choices might be going to just the TRI or FRS number for all databases.**

   We have identified the need to address not only the facility ID, but the different definitions of facilities used across air emissions reporting programs. For the facility IDs, we first need to determine how the facility definitions will be handled so that we can decide what a single ID would represent. This has not been decided, however, a new data model being developed by FRS seems like a promising option.

   Valid reasons can be found for the different facility definitions across air emissions programs, and we are not yet certain that forcing everyone into the same definition is needed or beneficial. In some cases, the facility definitions are a part of the Clean Air Act (TRI) and regulatory definitions, both of which are difficult to change. In other cases, the concept of facility stems from a certain regulation or agency only needing to know about certain emissions processes for that regulation. Therefore, we intend for the Shared Facility Attributes to be able to support the different definitions of facility, but unify those using some approach that allows for understanding what part(s) of a site are considered a "facility" for each regulation.

8. **Will there still be any redundancy at all regarding quality control/assessment? Sometimes looking at things twice catches major errors.**

   The overarching idea with the Lean approach is efficiency, which is intended to arrive at the appropriate quality for the least effort. As the QA aspects of the design are developed, if we determine that it is most efficient to duplicate certain checks (because it is the most efficient way to catch errors), then that approach could be selected. The proposed "to be" state would use as much automated QA as possible, allowing SLT agencies and the EPA to focus QA time and effort on the areas that add the most value to the final emissions estimates.

9. **The Clean Air Act of 1990 requires states to fund their programs through emission fees. If a company reports directly to EPA, then will EPA complete the environmental fees for each state?**

   Recall that this project is an exercise in joint governance among SLT agencies and the EPA. For this project, the EPA and SLT agencies are identifying data that the participants need and then are working collaboratively to determine the most efficient and effective ways to get accurate data.

   Since the emissions data would go both to the EPA and SLTs at the same time, SLTs would still be able to use their emissions to assess emission fees. Some SLTs have mentioned that the QA surrounding the fee assessment process helps to identify data problems. It will be important to ensure the same or better accuracy in emissions as part of the "to be" solution to support fees as one of the many uses of the emissions data.

10. **How will you include different lists of pollutants? For example, TRI has different pollutant list than NEI?**

    A solution could be to have a master list of pollutants that can be reported, but to allow each receiving air emissions program to select which pollutants would be submitted. In some cases, we

may want to harmonize the pollutant definitions (e.g., mass of metal/HAP part of a compound versus mass of entire compound) to streamline reporting.

11. **Presently TRI is doing and effort for between TRI and NEI and see big discrepancies, and data emissions.**

    This is the sort of difference that we would not observe in the future. If TRI or NEI emissions were updated after the data had already been submitted, both TRI and NEI programs would receive the updated data. Currently, since the reporting is completely separate, one program can receive updated data without the other program being informed (admittedly, an inefficient way to solve this problem is for manual communication of information, but this has not been practical to date). The "to be" solution would notify all parties automatically to allow for efficient awareness of the new data by all affected programs.

12. **I'm not sure the Turbo Tax® example of "state portals/systems" would address your concerns and the point of the lean project. The state system has its own key data (attributes).**

    Turbo Tax is just an example, and many details still must be worked out. The solution must meet the needs of SLTs to be successful. Fortunately, we have had great collaboration from SLT agencies on this project so that we are aware of these types of challenges. We intend that this project will incorporate state data systems where those data systems meet the requirements of the "to be" state. Updates to state data systems may need to be made for this to work. As part of implementation, the project will need to define the criteria that all data systems must would need to meet to become part of the "to be" solution. The facilities in each state would only benefit from the reduced workload of the "to be" solution if the SLT agency participates.

    One element of the Turbo Tax model that could work is the model for developing the submission software. Like the IRS specifying the tax submission requirements, the EPA and SLT agencies could specify the emissions submission requirements and third-parties could develop the submission software to meet those requirements.

13. **Will this effort accommodate or dovetail with the rulemaking efforts to require e-reporting and notification under air and other rules (March 20 proposal for NSPS will be followed by similar rule changes for MACT, NESHAP and Title V).**

    Yes, this effort is already dovetailed with the rulemaking efforts to require e-reporting of compliance test data. These will be the data that go to the CEDRI system, which is a part of our "to be" solution.

14. **How could the problems be solved in another manner, without developing all these systems and the need for ref and statutory changes?**

    With the Lean approach, we did not presume a technological change was required. We specified our problems as listed above in question 1 and then sought out root causes to those problems. Most of the root causes relate to our systems getting out of sync with each other, and so most of the solutions deal with keeping those systems in sync with each other. With some of the issues being extremely complicated (such as having the same facility defined in 3-4 different ways), it is presumably possible to manually resolve all of the related data challenges, but we expect that having something implemented systematically and through system-based data management solutions will be much more efficient. In addition to the systems, the "to be" solution must define

ways that the various parties can work better together. In fact, the way we work together is just as important, if not more important, than the systems that help us do that.

15. **In the Lean process, what feedback did you receive from our customers (industry) on their compliance status if they fail to submit reports electronically?**

    One of the major themes from industry about air emissions reporting was that there are so many separate reporting requirements, it leads to a lot of opportunities for missed reporting and therefore they can become out of compliance with reporting requirement. Our industry participants helped to identify a single way of reporting as a way to have fewer opportunities to miss compliance deadlines and have better compliance with reporting obligations overall.

16. **Who gets the final say on the "configuration" of a facility?**

    We are not far enough along in our project to have defined the business rules needed to answer your question. We have identified the need to make sure that all uses of the facility configuration are being met.

17. **When do you plan on going to the e-Enterprise Leadership team to see if you get the go ahead on this?**

    Since this webinar occurred, we have received approval to start a limited set of “short term wins” during 2015 and 2016 with total funding of $200,000. We have identified 6 projects that will do small steps along the path to the “to be” solution, but will have benefits as stand-alone efforts regardless of whether additional funding is provided for a more complete approach in the future. More information will be provided about these projects in the near future. States are invited to participate on those project teams.

18. **How are you costing out the downsides to this initiative for the ROI?**

    The ROI includes both the costs and benefits of the “to be” state.

19. **Will the E-Enterprise cover retrospective emissions information? For example, people could retrieve the historical emissions for a facility interested?**

    We intend to start with new data being reported so that our data partners can start realizing cost savings and data can be released earlier. Going back and refitting older data into a new model would have its own sets of costs and benefits and is not planned at this time.

20. **Will OEI be the lead on putting the system together?**

    As the lead on the Central Data Exchange, we anticipate that OEI will be a large part of the final development team, however, the roles have not yet been fully defined. As mentioned, under a Turbo Tax approach (see question 12), it’s possible that a third party could develop some of the systems.

21. How will the data from continuous monitors will be handled?

There are two uses of continuous emissions monitored (CEM) data. First, CEM data are reported continuously by facilities to the EPA through EPA's various air emissions trading programs. We have identified this data flow as one to be considered in the future, but have not explicitly included it in our analysis to date. Second, CEM data are collated for a year and reported as annual total emissions values by SLT agencies to the EPA for the NEI and in some cases (e.g., mercury) for the TRI. This latter data flow has been captured in our plans thus far. Since the CEM data are often needed as an annual total for meeting state needs (e.g., compliance and fees), we are assuming at this point that the annual data flow approach is sufficient for addressing the problems being solved by this project.

22. Who would be responsible for coordinating with reporters to make corrections to their submittals?

We are not far enough along in our project to have defined the business rules needed to answer your question. We anticipate that states would continue to be involved in this part of emissions reporting workflow and quality assurance, but there could be others involved as well.

23. How do we join the E-Enterprise Council?

The E-Enterprise Leadership Council has 20 members 10 states and 10 EPA. The bylaws for the E-Enterprise Leadership Council dictate the procedure for adding or changing members. The by-laws can be found in the E-Enterprise Blueprint at http://www.exchangenetwork.net/e-enterprise/.

24. You mention that a benefit would be to eliminate the need to augment NEI data. We find TRI data questionable and do a more robust HAP inventory in our state. How do you address data source preferences?

We already hope that we're receiving your state's HAP data even though it's voluntary. In cases where we are not receiving a given pollutant for a facility, but that facility has already reported emissions to TRI, we are using the TRI data (which is a facility total). We would prefer to get the detailed process-level data where it is available.

We learned in our Lean event that when some facilities estimate emissions to send to TRI, they do so at the process level and then add up the emissions to the facility level for TRI reporting. Since the NEI program and the TRI program both want the "best available" emissions, there is no "real" reason that the data coming to the NEI through the state at the process level should be different (when summed) from the facility TRI data. This project seeks to understand what reasons exist for such differences and identify streamlined ways to eliminate such differences. For example, no one is checking now to make sure that the same methods being used for state/NEI are also being used for TRI. The "to be" state would allow for such checks and also allow for data revisions to be sent to all parties at the same time so that emissions stay in sync.

To answer your question more directly, we anticipate that for the NEI, we would no longer need to do as much (if any) choosing one data source over another. To the extent that we will need to continue to do that as part of our "to-be" state, we do not yet know how we will do this. At the

current time, we use the state-reported emissions (often from the facility, but not always) preferentially.

25. I have found that TRI data is better data than state data....so disagree on this.

Sometimes it's hard to know which data source is better than another. However, since the NEI, SLT air agencies and TRI all want the "best available" HAP emissions estimates, these types of disagreements are greatly reduced if not eliminated through our proposed "to be" solution.

26. Can you identify which regulated company or organizations were worked with on the effort to date?

Alcoa, Phillips-66, and the Air Force. The latter provided an example of an organization that collects information from all facilities and reports simultaneously, while the industrial representatives provided a data flow directly from facilities to air agencies. We have also had input from members of NEDA-CAP, the Air Permitting Forum, and NCASI.

27. How dynamic do you anticipate the model to be for facilities that change operations year to year?

The model must be as dynamic as is needed to meet all of the needs of air programs. As part of the design process, it will be necessary to gather information on the needs by all air programs. Some of the features that have already come up include the ability to make changes that are date-specific, so that everyone can see when a change has been made and therefore determine the configuration of the facility at a given point in time.

28. How is this activity related to or incorporating strategies from John Dombrowski's work on what I would call "new data sources?"

The work of John Dombrowski's division complements this project. In that work, the EPA Office and Enforcement and Compliance Assurance (OECA) has developed a tool that displays emissions and other data from numerous sources in a single interface. That project is already aware of the efforts at EPA for a new data model for facility attributes, which this Combined Air Emissions project has also identified as a need. As the E-Enterprise work moves forward and is able to successfully implement the revised facility attributes approach, OECA's project and tools will benefit from those improvements in facility information consistency and removing the needs for facility matching. When Combined Air Emissions eventually provides improved and more consistent and timely emissions data, the OECA tools would also benefit from those improvements.

29. Will regulatory changes be an issue to move forward?

We believe that a great many things can be done without regulatory changes as we work towards our "to be" solution. We have a rough 5-year plan for this project and had anticipated that any regulatory changes would be made later in that time period. We would make regulatory changes only after we have learned more about the best way to proceed and have attempted to work within our existing regulations.

However, we are currently considering the various challenges associated with future changes to regulations that may be needed for this project. So, we know it will be an issue, but we haven't yet addressed that issue.

30. How will EPA avoid doing a rule when the different data systems have different submittal dates. In order to make this work, won't all the programs need the same submittal time?

We believe that it is not necessary to change submittal dates to keep the data in sync. For example, TRI has the first reporting deadline. After that deadline, the facility may also need to report HAP or ammonia emissions for a SLT regulation. If the reported emissions for that second purpose are different (a different emissions value) from the TRI reported emissions, a few things could happen. First, the reporting facility would see the discrepancy as part of the pre-submission QA for the SLT regulation. If the discrepancy is real, then the new emissions value could be sent to the TRI program as a revision so that the emissions value could stay in sync.

In addition, we have identified a need for changes that happen after the initial submissions as part of QA. Any emissions changes resulting from post-submission QA would be sent to all programs that had previous received the data, so that the new emissions values could be used and would be able to stay in sync across multiple programs.

31. Are you proposing to amend the AERR to eliminate the reporting to states and institute direct reporting to EPA?

The AERR does not require facilities to report to the state. The AERR requires states to report to the EPA, and we do not expect to eliminate that requirement. We have not yet determined what regulatory changes may be needed to accomplish the full “to be” solution, but we expect to learn a lot more about those needs by starting with approaches that do not require changes to regulations.

32. What is the outreach plan to get a majority of states directly involved in this planning process?

The Scoping Team is in the processes of drafting an outreach and communications plan as part of the more detailed project planning that will happen in the summer and fall of 2015.