

## **Combined Air Emissions Reporting (CAER) Frequently Asked Questions**

### **E-ENTERPRISE**

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

The e-Enterprise Leadership team has given us the go ahead to proceed, but we are awaiting the majority of the resources we expect that we will need for this project. In fall of 2015, this project received approval to start a limited set of “short term wins” during 2015 and 2016 with total funding of \$300,000. We have been working on 5 projects that are taking small steps along the path to the proposed future solution, and these short term wins will also have benefits as stand-alone efforts regardless of whether additional funding is provided for a more complete approach in the future.

In FY 2016, we have requested additional resources to continue more short term wins that have proposed to test the use of the new FRS data model for supporting the central and detailed facility attribute needs for this project.

The FY 2017 President’s budget includes a line item for this project funded at more than \$2 million, and so we are optimistic that at least some resources will become available by 2017.

**2. 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.

**3. 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 bylaws can be found in the E-Enterprise Blueprint at <http://www.exchangenetwork.net/e-enterprise/>.

### **OVERALL COMBINED AIR EMISSIONS REPORTING PROJECT**

**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 7 from the January 2016 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 prolonging 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 using best available data 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 not 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.**

The Combined Air Emissions Reporting (CAER) project seeks to streamline multiple emissions reporting processes and to establish a single, authoritative data repository. Currently, air emissions information is collected by EPA and state or local air agencies through numerous separate regulations, in a variety of formats, according to different reporting schedules, and using multiple routes of data transfer. In the to-be state, the CAER project is expected to reduce the cost to industry and government for providing and managing important environmental data and to improve decision-making capacity through more timely availability of data. The proposed to-be state for the air emissions programs is intended to reduce the cost to industry and government, or "transaction costs," of providing and managing important environmental data. Ultimately, the improvements in the to-be state might increase the use of CAER information by government and non-government entities, potentially leading to more informed decisions and better outcomes. Please see the answer to question 1 above.

The Return on Investment (ROI) analysis shows that the proposed to-be state for the CAER project might yield reductions in burden for companies submitting air emissions data, as well as the agencies that process and manage these data. Overall, the ROI analysis indicates positive but uncertain savings. It is

important to note that the scoping team did not seek to reduce the level of information that is collected and made available to government and non-government data users. The goal of the project is instead to minimize the transaction costs of collecting and managing the data.

**3. 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.

**4. NEI inefficiency: Are you referring just to the point source category or the whole process? How much are nonpoint sources and categories other than large-point sources planned to be integrated?**

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. There have been questions about the NEI, which also includes "nonpoint sources", which are summed up at the county level and not characterized in any inventory as an individual source at a location or address. The goal is that the universe of facilities that will be included will be the macro set of all the facilities included in at least one of the four programs. Initially, though, we might focus on facilities that are in two or more of the programs, because that's where more benefit will be derived. We will go through a process to identify the most important things to do first, that have the greatest return on investment.

**5. Have you considered starting with two or three databases as a pilot?**

CAER teams are taking a stepwise approach and only starting with a small set of databases to pilot any potential changes.

**6. 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. Participants in this project have noted that one aspect of the facility configuration is the permits, and so our work will include making sure that the electronic data about the facility respect the permits and the state role in quality assurance and approval for the vast majority of these.

**7. 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.

**8. Who is 18F?**

18F is an incubator group at the General Services Administration (GSA) bringing in modern techniques to advance programs. They are helping E-Enterprise break down projects into bite-sized pieces so that we can come up with meaningful and useful components quickly in order to start producing solutions rapidly instead of waiting for the delivery of a massive system that may be out of date once developed. For more information please see: <https://18f.gsa.gov/>.

## **COST/BENEFITS**

- 1. How are you costing out the downsides to this initiative for the ROI? How did you quantify the additional cost to state/local/tribal (SLT) agencies for having to develop new systems, API's, web services, or other connection protocols to transmit data in the system or between distributed databases?**

The ROI includes both the costs and benefits of the proposed future state. The scoping team received input from many states participating in the project and Arizona issued a survey to states to quantify their existing costs. We didn't get an exact cost estimate from every state; it's difficult to develop cost estimates when we don't know exactly what changes will be needed. This made conducting a return on investment analysis challenging and has resulted in a lot of uncertainty. But there was a positive range of potential benefits to these types of changes, even including that uncertainty.

## **REGULATIONS**

- 1. 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 SPPD rulemaking efforts to require e-reporting of compliance information (e.g., performance tests, notice of compliance status, and air emission reports). These will be the data that go to the CEDRI/WebFIRE system, which is a part of our proposed future solution.

- 2. 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 future 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.

- 3. Will regulatory changes be an issue to move forward? Have you identified all of the rules in effect that may need revision make this work?**

We believe that a great many things can be done without regulatory changes as we work towards our proposed future state 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, CAER teams are working on identifying rules changes as they continue to work on the data flows. 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. One of the steps that we are taking now is to compile all of the references to the state regulations and capture each definition of "facility" and the reporting requirements for air emissions.

**4. 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 previously received the data, so that the new emissions values could be used and would be able to stay in sync across multiple programs.

**5. 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 proposed future state solution, but we expect to learn a lot more about those needs by starting with approaches that do not require changes to regulations.

## **QUALITY ASSURANCE**

**1. How will shared roles for quality assurance/control (QA/QC) work in situations where states QA/QC every item and every entry from their point sources before submitting to the EPA? 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. We have to ensure that the CAER proposed future state respects and allows

for the EPA and the states to retain their QA/QC programs. It is critical that data are not used before going through the proper vetting process. In addition, QA/QC steps should occur as early in the data exchange as possible. The CAER teams are discussing QA/QC and are beginning to identify places where more clarity or additional review could help identify data issues or potential outliers for documentation. Communication is key to minimize the number of times data must be reviewed and resubmitted between the states and EPA. 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 future 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.

**2. Are you considering the data that you're getting draft until such time that the state says they're done with their QA/QC?**

We do not yet know if it will be considered draft, how it will be treated, when EPA will get it, etc. These are business rules that need to be resolved so we can proceed with implementation. We must take precautions to ensure the data doesn't get utilized improperly or in a misleading way if it hasn't been through the QA/QC process. It's unclear how we will ensure that QA/QC occurs, at what point we do editing, or who has access to make those corrections. We want industry to help inform states and EPA as we work together to create a CAER future state.

**INDUSTRY**

**1. 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.

**2. What is the workload for industry and how might that change? Will it spike because things have to be done at the same time?**

The feedback from our initial focus groups and our lean event is that the industry currently appreciates that the deadlines are spread out. Right now the GHG Reporting Program and TRI deadlines are earlier in the year than the (some) state deadlines for NEI; this allows industry staff to be more efficient and effective. Some of the benefits to industry and agencies from this system are related to the linking of the facility information. If a piece of information comes in that's inconsistent with data that was submitted earlier on, it can be flagged for the affected parties both at the facility as they're submitting it, and, if necessary, at the various agencies. Agencies will be notified of discrepancies, and they can decide how to proceed.

**3. 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, AF & PA, and NCASI.

**4. From an industry perspective, creating an electronic air emission reporting tool that States could adopt and use for their annual reporting requirements would be a great improvement over the current reporting process. Is this part of the CAER initiative?**

Generally the answer is “yes” in terms of a possible conceptual design, but will depend ultimately on the development details especially as it involves interface with individual state reporting systems. For example, delegated programs like NAAQS emissions reporting would need to include the state leadership roles in whatever technology solutions are ultimately created. This would allow the states to continue to have the lead on collecting that information such as the emissions data for the National Emissions Inventory program. This is especially important for the states because the fees generated from the emissions reports are a part of their state program funding.

**5. How could activity be CBI for some programs but not others? For example, where might it affect one program and not another?**

In certain contexts programs have responded to a claim of CBI without addressing the legal question of whether it's CBI or not, and the program has responded by treating it as CBI. Those decisions have not been made consistently across all the states or EPA programs. If different programs have a disagreement about what's CBI, we need to address those discrepancies which would otherwise interfere with consolidation.

**6. How dynamic do you anticipate the facility attributes data 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

**7. When might industry see first changes? What's the timeline?**

The implementation team (short term win project) is working on development of the schedule as part of a high level implementation plan. This is a multi-year project, where we're targeting 5 years but working incrementally such that some functionality could be available sooner if we can work out the details. The goal is to make significant progress within five years, but we want to see some early wins and make improvements sooner than that. We don't know when that might start affecting individual facilities. The high level implementation plan is expected to be available later in 2016.

**STATE/LOCAL/TRIBAL**

**1. 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 proposed future state solution without “starting over”.

In addition, we would like to use an approach moving forward that first seeks to help those states that need the most help. For example, some states do not have electronic reporting systems or are not happy with the systems that they have. Through this project, we can focus on those states that need the most help first, while keeping in mind that some states who have robust systems would likely need to tie into the CAER approach in a different way such that their system is still largely intact.

**2. Are State emission inventory reporting requirements different from EPA requirements? How will these differences be accommodated?**

We want what we're collecting for state purposes to align with what we're collecting for EPA purposes. We are aware that states collect data that goes beyond what is required for EPA programs. In some cases, states are voluntarily sharing that information, while in other cases, they do not share it. For States that are willing to share all of the data they collect, CAER could provide a way for them to more easily do that.

One concern has been raised that when states are not able to share a piece of data, how would that work without undue burden on the state? We want states to be able to make choices about data that are not required to be submitted; therefore, this will be a key consideration for the design and implementation phase of the project. If there are cases where the state's requirements are less than EPA's, yet the state is still required to report to EPA, then CAER can help to identify and ultimately solve those problems.

**3. If states require things like ozone season data, are we going to lose that data because the EPA isn't going to ask for that?**

No. We wouldn't recommend removing things just because they're not required by EPA. The goal is that all program needs are met and that the reporting is streamlined.

**4. Would you ask for a facility in New Jersey to submit something that another state might not ask for?**

Yes, but the EPA wouldn't necessarily be the ones asking for it. Each of the states' needs must be met for their various reporting requirements. Those regulations are coming through the states because for the NAAQS, the EPA has largely delegated the program to the states. We assume that will continue to be the case. We hope to avoid repetitive reporting. If a certain state needs more data, this system would hopefully provide the opportunity for that data to continue to be collected as it has been previously. Only the data required for EPA would feed into the appropriate systems and programs; the goal is to not make a current state system unworkable. This process will not prevent states from collecting necessary

data, nor will it require a state to collect data just because a different state is collecting it. CAER will recognize and respect that states have individual needs for data. The challenge will be making this seamless and maintaining the necessary data protections, while granting states access to data.

**5. Will state systems have to speak "live" to federal systems during reporting periods while industry is entering data?**

The concept of live communication among systems is yet to be determined, but it's certainly possible. If the facility attributes are transmitted to the federal database before the company has certified the data, this demonstrates the need to determine business rules. The goal is for only quality information to be shared, and we want to ensure efficient data processing, uploading, transmitting, and vetting. Many facility attributes are tied to the permits, and so there needs to be some ability to ground-truth the electronic data with the permit. We will have to address whether multiple people can be making changes simultaneously, who has the authority to make change, who and how can changes be quality assured and approved, and how we will track changes over time.

**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 proposed future 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. When CAER allows for easy sharing of data from 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 CAER "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 future 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 other systems in the proposed future state. The advantage of such an approach would be a mechanism by which submitted data can be more easily submitted, quality assured, and shared amongst the various reporting programs.

**7. A previous audit of the Superfunds Amendments and Reauthorization Act (SARA) Title III section 313 data compared it to a state emissions inventory data (facilities that possibly should be reporting but are not). This audit revealed differences in how facilities are reported (e.g., as groups vs. separate instances). Is there a proposed solution for reconciling facility reporting differences across the databases?**

The CAER team is working on this issue and collaborating with an existing E-Enterprise team focused on federated facility ID. A pilot will be employed to identify where these differences exist and how they may be addressed. Historically, each facility reporting requirement contained an acting statutory requirement and each of those statutes had various thresholds for reporting. One benefit of the proposed approach is that the different facility definitions would become much more transparent through the shared data approach. So, audits of the types that you mention could theoretically be run as a simple report rather than require a lot more effort.

**8. How will the CAER future state affect existing state emissions reporting systems like SLEIS? Will sources need to report NEI emissions to SLEIS or directly to CAER?**

SLEIS is an existing vendor-developed system that many states have adapted for their own needs. CAER is investigating communication options using SLEIS to other systems and portals to leverage existing infrastructure. It is possible that a state emissions reporting system such as SLEIS will be able to export the data directly into the CAER system. One goal is to maintain functioning data reporting systems maintained by state jurisdictions where that is desired.

**9. On the Data Flow diagram (slide 8 in the January 2016 webinar), facilities only report into the main CAER portal, not any state portals. Is that what is envisioned for this project?**

Ideally, users will be able to continue using their state portal with data exported directly into the future CAER central collection in a standard data format. This diagram should be considered from the facility's perspective, where the desire is to have a unified experience for the facility. A single portal for the facility does not necessarily mean that there is a "main CAER portal" used by every facility, but rather a facility could only work within a state portal that has been enhanced to support other programs as well.

Example (may not apply to all facilities): If a facility or company is currently reporting to Texas, but Texas doesn't have oversight of TRI, how could their experience be unified across all programs? Technology solutions may allow Texas and TRI reporting approaches to be connected. Ideally for the facility, everything would be unified through a single Texas portal, but such a portal is not necessarily the same portal that a different facility in a different state would use.

**10. 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. While the data may go to all parties at the same time, it would not be certified for some

or any uses until the appropriate QA/QC has been done commensurate with requirements associated with fee assessment.

**11. 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 proposed future 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 would need to meet to become part of the solution. The facilities in each state would only benefit from the reduced workload of the future 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.

**12. Do you expect a minimum number of states to join in before you move forward with implementing the enterprise system?**

Because of the stepwise nature of what we're doing, we don't necessarily need to have a minimum number of states. In some ways it would be advantageous to start with a very small number who were willing to experiment and work through some of the challenges early on, so that other states can benefit from that knowledge. If we're able to get this to work for even one state, all of the industries in that state would benefit and then others would be even more inclined to join. The more states that are involved, the more costs are spread, and the more value will be received for the money and time invested. However, when there are fewer states starting initially it is easier to manage the challenges encountered. If we have fewer challenges and issues, there will be more people wanting to participate. If it's more complex, there will be fewer willing to participate initially. Long-term we want this to remain voluntary, where people participate when it adds value.

**EPA**

**1. 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. We do not believe that existing facility definitions across programs would need to change, but that the central storage of a facility needs some unifying concept around which the other facility definitions can build. A

new data model being developed by FRS seems like a promising option, which we are currently exploring as part of the CAER effort.

Valid reasons can be found for the different facility definitions across air emissions programs, and we do not believe 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 needing to know about only 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.

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

These implementation details have not yet been determined. One 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. The issue of sharing codes tables is a general issue that affects not only pollutants, but other codes that are central to the sharing of information.

**3. Presently TRI is doing a comparison 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 CAER is successful. 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. While we could be (but are not) manually communicating these differences now, such an approach has not been practical to date). In the proposed future state, the system would notify all parties automatically to allow for efficient awareness of the new data by all affected programs.

**4. Not all emission source types have emission factors in WebFire or AP 42. Is there a way for users to post emission factors from source-specific stack testing or CEM data to be used for calculating emissions?**

The Compliance and Emissions Data Reporting Interface (CEDRI) deals with source test data collection and includes other components such as the emissions reporting tool, ERT, and WebFIRE. The goal is for these systems to collect mandatory compliance source testing information. Voluntary submission of other available test data are also encouraged to be provided using these tools, as long as the source testing parties follow the methodologies that need to be used for doing so. Additional test methods must be programmed for both ERT and CEDRI, and so one caveat here is that the tools must support the test method being used for stack testing. Through CAER and CEDRI, the EPA wants to compile all available source test data and periodically use it to assess whether emission factors need to be revised with in WebFIRE. When emission factors do need revision, the EPA would use the notice and comment approach to propose and finalize new emission factors that has been used in the past.

- 5. Of the four reporting programs included in E-Enterprise, the information collected in CEDRI appears to be fundamentally different than the others. It is short term and based on an individual emission source. The data collected for the balance of the programs are longer term and may be facility wide. What is the benefit of collecting the CEDRI data as part of this effort?**

As described above, the ERT, CEDRI, and WebFIRE systems together are forming a repository of compliance test data and a renewed basis of data for assessing and computing revised emission factors. In addition, CEDRI and the reporting facilities also benefit from the sharing of facility attributes so that information does not need to be resubmitted to the EPA (it can be prepopulated from the Shared Facility Attributes part of the proposed future state). Sharing that common understanding of a facility across all programs will have enormous benefits and resource savings. For example, if the test reports are connected to a common understanding of the facility that includes previous test data and emissions reports, the facilities and agencies may be able to more quickly assess when a problem exists in emissions, more quickly correct any problems, and more quickly compile that information for communicating with the public. Having all of these pieces integrated will provide us with more opportunities to leverage that type of information for various purposes over time.

- 6. Is ICIS-Air going to be included in this project?**

ICIS-Air was not included in the first round of four programs to investigate. However, discussions are ongoing now within the EPA to make sure that the Shared Facility Attributes and other key pieces of the proposed future state will be beneficial to the ICIS-Air system. The ICIS-Air program is a user of data coming in through CEDRI and the inventory collection, and so we will be looking at how best to connect CAER with these types of users as we make progress on this project.

- 7. How is this activity related to or incorporating strategies from OEI's work on "new data sources?"**

This question relates to the previous question. The work of OEI complements this project. In that work, the EPA Office of 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 CAER 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 CAER eventually provides improved and more consistent and timely emissions data, the OECA tools would also benefit from those improvements.

- 8. Will OEI be the lead on putting the system together? How does this relate to the central data exchange (CDX)?**

As the lead on the Central Data Exchange (CDX), we anticipate that OEI will be a large part of the final development team, however, the roles have not yet been fully defined. The CDX is a piece of technology for reporting to EPA that is currently used by many of the emissions reporting programs. Updates to the combined air emissions reporting plan to leverage this database and the shared services

already developed by CDX. As mentioned, under a Turbo Tax approach (see question 21), it's also possible that a third party could develop some of the systems.

**9. 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.

We are also considering whether the CEM reporting program electronic reporting can benefit from the Shared Facility Attributes concept envisioned in the CAER future state.

**10. 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 may be 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 CAER future 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 future 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.

**11. I have found that TRI data is better data than state data....so disagree on the previous comment.**

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 future state solution.

## **OUTREACH**

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

We have worked with the Environmental Council of States (ECOS), NACAA, and AAPCA thus far. We have briefed the Air Program Managers and Air Directors at the states. We have requested state participation in this project many times since the fall of the 2014. Every state has therefore been contacted on many occasions and at many levels to be involved.

We have numerous states involved at this time as shown in the webinar slides. There will continue to be new opportunities for states to participate in upcoming work as additional facets of the project move forward. All states can have a voice in this process using the coordination mechanisms offered by ECOS, NACAA, and AAPCA.