

PRTR



The Challenge of Collecting and Maintaining Quality Data – Overview of U.S. Toxics Release Inventory Data Quality Activities

*U.S. PRTR –
Toxics Release Inventory (TRI)
December 2007*



Overview of U.S. TRI Data Quality Activities

Data quality efforts BEFORE facilities submit data:

- Availability of TRI – Made Easy software
 - Interactive, user-friendly software that guides facilities through the TRI reporting experience
 - Intelligent reporting software that includes many built-in data quality checks
 - Allows facilities to report via diskette or over the Internet
 - Moving from PC-based software to web-based software (i.e., TRI-ME web) that will include enhanced data quality and validation assistance
- Availability of Reporting Forms and Instructions Package
 - This document provides a step-by-step explanation of how to complete the TRI forms and contains a wealth of information designed to minimize reporting errors.
 - The document is updated annually and made available through the TRI website.



Overview of U.S. TRI Data Quality Activities

- Availability of technical guidance documents
 - The program has developed no less than 30 guidance documents to assist facilities in estimating and reporting their release and other waste management information
 - Both industry-specific and chemical-specific guidance documents are available
- Training Workshops
 - In-person workshops throughout the country
 - Also developing training materials for use over the Internet
- Frequent Questions Service
 - A service available from the TRI website which makes it easy to obtain answers to frequently asked questions and to submit new questions about TRI.
- TRI Information Center
 - A toll-free number that industry representatives may call to obtain direct answers and guidance to specific questions about completing TRI Forms



Overview of U.S. TRI Data Quality Activities

Data quality efforts AFTER facilities submit data:

- Data Quality efforts at the TRI Data Processing Center (DPC)
 - The DPC receives all TRI data and runs a number of data validation and data quality checks
 - As much as possible, data quality checks done at the DPC are being moved into the “front-end” of the reporting process via TRI-ME web
- Electronic Facility Data Profile (eFDP)
 - Facilities that submit TRI data to EPA are notified that they have an eFDP posted on a secure website
 - The eFDP “echoes back” the submitted data and presents informational messages about the validity and quality of the data as submitted by the facilities
 - As much as possible, the eFDP data quality checks are also being moved into the “front-end” of the reporting process via TRI-ME web



Overview of U.S. TRI Data Quality Activities

- TRI In-house Data Quality Analysis
 - Each year, months before EPA makes available its latest annual Public Data Release, the TRI Program conducts an in-house data quality analysis on the data about to be released.
 - Algorithms are employed to screen the data for outliers, and reports are generated identifying facilities that have shown substantial changes in the quantities of releases they've reported for TRI chemicals relative what they reported in the previous year.
 - Another report is generated that identifies an extensive list of anomalies in the data. All these reports are reviewed by staff and a list of facilities that have what appear to be probable reporting errors is formulated.
 - Technical staff within the TRI Program contact these facilities to discuss their submissions and, where errors are confirmed, have the data revised.



Overview of U.S. TRI Data Quality Activities

- State Reconciliation
 - EPA generates a series of reports listing all the facilities and chemicals that have been reported from each state. The reports are sent to each state DEP. The states review the reports and determine which data, if any, they have collected that EPA has not and vice-versa.
 - Facilities that have neglected to send in data are then contacted.
- Contractor Review of the Data prior to an Annual Public Data Release (PDR)
 - EPA conducts a number of tests looking for data anomalies. When suspicious data are discovered, more focused analyses and research are performed to determine if the anomaly is an error or not. Facilities are contacted, questioned and, where necessary, are told to revise their submissions.
 - This is the final phase of data quality efforts before data are made available to the public.



Overview of U.S. TRI Data Quality Activities

- Cross-Referencing Project
 - EPA's TRI Program routinely compares TRI data with similar data contained in other EPA databases [e.g., National Emissions Inventory (NEI), Permits Compliance System (PCS) and Biennial Reporting of Hazardous Wastes (BR)] to identify possible non-compliance and/or data quality issues.
- Error Tracker and TRI-Explorer inquires
 - Envirofacts and TRI-Explorer are the two main data access tools used by the public.
 - Both systems offer options for data users (e.g., citizens) to report possible errors. These reports are provided to TRI staff for review and possible action.



Overview of U.S. TRI Data Quality Activities

- EPA Regional Data Quality Activities – EPA’s Regional Offices play a critical role in optimizing the quality of TRI data. They undertake a variety of activities that help to maximize TRI data quality, these include:
 - annual TRI training workshops;
 - offer direct and on-demand compliance assistance to facilities (e.g., regularly discuss complicated TRI issues with facilities, have face-to-face meetings with facility personnel);
 - contact facilities who’s most recent Form R reports contains substantially different information from that submitted the previous year;
 - collaborate with staff from other federal EPA programs or state EPA programs that also collect waste management data on chemicals included on the TRI list of Toxic Chemicals, to identify and resolve any discrepancies in information submitted to the TRI Program and other EPA programs (federal or state);
 - call or conduct on-site visits to facilities for which citizens or state environmental staffers have raised concerns about reported waste management data, or who appear to be submitting data that are erroneous particularly those facilities with release reports that show an unusual upward or downward trend;
 - conduct chemical-specific (e.g., PBT chemicals) or industry-specific targeted reviews of Form R release reports to identify possible reporting errors;
 - provide compliance assistance through e-mail campaigns or mailing of flyers (e.g., such as when changes to the TRI reporting requirements occur, or an a new guidance document has become available)



Overview of U.S. TRI Data Quality Activities

Some Possible Future Data Quality Activities include the following:

- Use of More Accurate Emission Factors
 - Facilities in certain industry sectors appear to be using emission factor values to estimate release quantities that are not as accurate as some newer emission factor values.
 - The TRI Program would like to proactively encourage facilities to use newer emission factor values that are known to be more accurate than older emission factor values.
- Industry Sector Analysis
 - Generally, facilities in the same industry sector have similar processes, and use and produce the same or similar input and output materials, respectively.
 - Plans are to profile an industry sector (e.g., electric generating facilities and petroleum refining for 2008) to better characterize the types of emissions facilities within the sector. Armed with this knowledge, EPA will be better able to identify potential reporting errors from facilities within the sector.



Overview of U.S. TRI Data Quality Activities

- TRI-ME web
 - Continued deployment of TRI-ME web (i.e., web-based reporting software) to include enhanced data quality and validation assistance
- Statistical-Based Identification of Potential Reporting Inaccuracies
 - With the advent of advanced computational power and statistical capabilities, the potential now exists to apply more sophisticated statistical data analysis and data pattern recognition methods to TRI data, which may provide more novel, targeted approaches to identifying unreasonable estimates of chemical releases.



For Further Information

Michelle Price
U.S. Environmental Protection Agency
Office of Environmental Information
Toxics Release Inventory Division (2844T)
1200 Pennsylvania Avenue, NW
Washington, DC 20460
phone: 202.566.0744
fax: 202.566.0741
Email: price.michelle@epa.gov
www.epa.gov/tri

