Regulator Seeding of Facility Data in CERS2

Prepared by Cal/EPA Technology Services Staff Draft Version 2 November 2011

This paper discusses the process Cal/EPA has established for CUPAs to "seed" facility data into CERS2 using specifically defined Microsoft Excel spreadsheets. Cal/EPA defines "data seeding" as an optional process **CUPAs** can use to perform a **one-time**, initial load of a limited subset of facility data. Once loaded, this data becomes one or more "draft" submittal elements for existing or new facilities in CERS2 which would be confirmed (and potentially modified) by a reporting entity prior to submitting the data to the regulator.

This paper is primarily targeted toward project managers and technical staff responsible for their regulator's CERS efforts. However, the bulk of the paper's content is generally of a process/procedural nature. The appendices discuss the Microsoft Excel seeding templates and field references, which provide the technical specifications the regulator's technology staff, contractor, or vendor will use for preparing seeding data.

Note about this document's "Final Draft" status:
The November 2011 version of this document reflects the reduction in the number of minimally-required fields for facility information shown in the October 2011 version. This draft is being released as a "final draft" to collect comments on any additions/modifications which would make the instructions clearer. However, the overall process/procedure will not be changing in the final version.

FINAL DRAFT

Table of Contents

Background

When CERS2 is deployed in winter 2011, it will contain approximately 10,000 facilities migrated from CERS1/UNIDOCS. The remainder of the approximately 140,000 Unified Program-regulated facilities will be added to CERS2, referred to as CERS in the remainder of this document, over time using one of the following mechanisms:

- Manual entry of facilities directly into CERS by facility owner/operators.
- One-time, batch addition of facilities by Regulators using the "data seeding."
- Electronic addition of new facilities using CERS EDT.

This paper provides the process and templates regulators will use for "data seeding" addition of facility data.

Please note that as regulators perform data seeding, the facility owner/operators might be adding or modifying facilities in the CERS Business portal at the same time. CERS does not preclude any regulated facility from manually entering and/or updating their facility information into CERS at any time. A number of large, multi-facility businesses covering multiple jurisdictions are and will be entering their facilities in CERS throughout 2012 independent of any actions performed by regulators. However, the majority of facilities are not owned/operated by multi-facility and multi-jurisdictional businesses/organizations, and will typically be following the guidance of their local regulators about when they should start using CERS (or a local portal). Many regulators will advise their regulated facilities to defer using CERS until after they have completed their data seeding (or to use their local web portal if one is being developed).

Once a CUPA completes seeding their facility data into CERS, they need to ensure any future data reported to CERS is either updated into their local data system(s) via EDT or manually updated. Failure to do so will cause inconsistencies between the regulator's local data system and CERS.

Overview of the Seeding Process

Below is a summary of key points about CERS Data Seeding

- Seeding is a means for regulators to move their facility data into CERS. Regulators might choose to seed their facility data in CERS to assist their facility owner/operators, or to ensure the accuracy of data in their local data systems, and/or to prepare their local data for CERS EDT efforts.
- CUPAs are NOT required to seed data in CERS. CUPAs with minimal or no **electronic** records of their facilities may wish to skip this process entirely, and instead instruct their facility owner/operators to directly enter their information in CERS. The regulator could then later download their facility information from CERS as necessary.
- Only a subset of the entire CERS data fields can be seeded. All regulators participating in data seeding must provide CERS Facility Information Submittal Element data (Owner/Operator and Business Activities forms). Regulators can also opt to seed chemical inventories and/or UST Form A/B data. Uploading supplemental documentation like site maps is not supported.
- Data seeding can only occur once per CUPA. If there are any problems with a submission record, the record will be rejected and returned to the regulator to resolve problems.

FINAL DRAFT

- CUPAs with PAs must coordinate efforts with their PAs and compile all data CUPA-wide into a single set of data for submittal to Cal/EPA.
- All facility data seeded into CERS becomes one or more "draft" submittal elements which would be confirmed (and potentially modified) in CERS by a reporting entity prior to submitting the data to the regulator.
- If a business/organization subsequently modifies the seeded data in a manner unacceptable to the regulator, the regulator should set submittal element's status to "Not Accepted" along with comments about how it should be corrected.
- CUPAs using local web portals are performing data seeding primarily to ensure any existing CERS facilities are added to the regulator's local data systems and to acquire a CERS ID for all of the regulator's facilities.
- The data quality/completeness requirements for data seeding are less stringent than for EDT or manual data entry. However, CUPAs must supply seeding data in the format proscribed in this paper or Cal/EPA will not accept the data.
- This paper includes a set of three data seeding templates in Microsoft Excel format that specify the format, minimal fields, and code values CUPAs must use for their data seeding.
- CUPAs are responsible for ensuring they have technical staff, contractors, or vendors with sufficient skills and time to prepare their data for seeding in the Cal/EPA specified formats. Cal/EPA cannot act as technical support to a CUPA for data seeding.
- A preliminary step in data seeding includes CUPAs and PAs downloading current CERS facility data and integrating this data into their local data system(s). This integration effort includes identifying duplicate facilities, identifying current data by comparing the CERS "last updated" field to an equivalent field in the regulator's data system, ensuring the most current data is moved into the regulator's data system in advance of the included seeding effort, and adding an appropriate CERS Organization Code for seeded facilities when possible.

The basic steps for seeding are summarized below. The remainder of this document expands upon this process.

1. Download listing of existing CERS facilities & CERS Organizations
2. Identify the CUPA's existing facilities in CERS
3. Reconcile data inconsistencies for CERS & CUPA facility records
4. Ensure CUPA data system can export CERS-compliant data
5. Identify CERS Organizations for facilities
6. Load CUPA facility data into seeding template spreadsheet(s)
7. Send facility data spreadsheet(s) to Cal/EPA
8. Spreadsheet(s) data loaded by Cal/EPA
9. Timely CUPA follow-up as directed by Cal/EPA
10. Repeat Steps 6-9 for corrections/additional elements as necessary

FINAL DRAFT

CERS Data Seeding Process

Preliminary Tasks:

- CUPA's Program Manager reviews and signs the CERS Usage Agreement (See Appendix C).
- CUPA's program and IT staff review this document ("Regulator Seeding of Facility Data in CERS2") and the accompanying MS Excel seeding templates (See Appendix A).
- The CUPA's CERS Manager or designee identifies any alternate contacts that will take the lead on data seeding efforts (i.e., information technology staff/contractors).
- Review the CUPA's location in the CERS Data Seeding Queue (see Appendix C). Seeding is anticipated to begin in February 2012 and should be completed by the end of 2012. Contact Dan Firth at dfirth@calepa.ca.gov if your CUPA's order in the queue needs to be adjusted.

Step 1: Download Listing of Existing CERS2 Facilities

One of the CUPA's authorized CERS users downloads the current statewide listing of facilities in CERS. This Microsoft Excel file contains the following:

- One worksheet provides all data fields making up the *Facility Information* Submittal Element for every facility in CERS as of the day before the download.
- A worksheet lists all CERS2 Organizations along with their CERS Organization Code, headquarters location, and facility count.
- A worksheet of all facility names, addresses, and CERS Organization Codes not already in CERS but reported by multi-facility organizations to Cal/EPA as part of their efforts to establish a CERS Organization (see https://cers.calepa.ca.gov/tempdocs/CERSOrgInfoRequest.pdf).
- A final worksheet including instructions and other information.

Information on how to download this file from CERS will be available at the CERS Data Seeding web site (see Appendix C). CUPAs needing an extensive amount of time to prepare their seeding data may need to download this file multiple times to maintain current data. The CUPA is encouraged to download this data and perform initial evaluation of the existing CERS facilities **well in advance** of the CUPA's seeding timeframe to better understand the scope of their efforts (i.e., how many of the CUPA's facilities already exist in CERS, when were these records last updated, etc.).

Step 2: Identify the CUPA's Existing Facilities in CERS

Review the statewide listing of facilities, and identify all facilities that belong to the CUPA (and its PAs if applicable). This can be done by filtering the spreadsheet on the *CUPA* field. However, consider alternate filters (e.g., by City or ZIP code) in case other facilities regulated by the CUPA are incorrectly assigned in CERS to a different CUPA. If an existing CERS facility is associated with an incorrect CUPA, confirm with the other CUPA what the correct regulator should be and update the association in CERS (or contact the CERS Help Desk at cers@calepa.ca.gov).

Step 3: Reconcile Data Inconsistencies for CERS & CUPA Facility Records

For the facilities identified in Step 2, review the CERS *UpdatedOn* field (the date/time the field was last saved/updated in CERS) and an equivalent field/indicator in the regulator's data system to determine which

FINAL DRAFT

system has the most current data. If the CERS record is more current, the CUPA should ensure all of the more current CERS data (especially the CERS ID and Organization Code) are imported into the regulator's data system **prior** to seeding. Failure to do this will result in duplicate files and/or conflicting data which will require future and more difficult reconciliation processes. Cal/EPA may also reject a CUPA's seeding files if they contain excessive duplicate records.

For CUPAs who have not been actively encouraging use of CERS1, there typically should be few existing CERS facilities to review/integrate. For CUPAs who have more actively promoted use of CERS1, there will be more records to review/integrate, but the majority of these records will already be associated with the regulator's previously assigned Facility ID.

If the CUPA identifies facility records which are obviously test, sample, duplicates (within CERS), or otherwise erroneous, the CUPA can either delete them in CERS, or send a spreadsheet to CERS Help Center with rows showing the CERS IDs, Facility ID, facility name, and reason for deletion.

Step 4: Ensure CUPA Data System Can Export CERS-Compliant Data

Regulators should review the CERS Data Seeding Spreadsheet Templates to ensure their local data systems (or IT staff/consultants) can export and/or transform data into CERS-compliant data. Issues to review include:

- Providing the minimally-required fields for each submittal element as indicated in the CERS Data Seeding Spreadsheet Templates (and in Appendix A)
- Returning appropriate codes values for various fields as described in the CERS Data Registry.

These two checks are especially critical. If the seeding data sent to Cal/EPA has missing minimally-required **data seeding** fields (which are less restrictive than "CERS Minimally-Required Fields") or invalid field types/codes, the spreadsheet will be rejected and returned to be corrected. Repeated seeding problems could result in the CUPA being significantly delayed as Cal/EPA moves on to process other CUPAs on the seeding schedule.

A more comprehensive checklist for reviewing the quality of regulator seeding data is shown at the end of this paper.

Step 5: Identify CERS Organizations for Facilities

To assist organizations/businesses with multiple facilities in accessing all of their CERS facility records with a single sign-in, the CUPA should review the facilities in CERS and their own system and identify any facilities which belong to multi-facility organizations. Determining which facilities belong to which organizations can be difficult, and during the seeding process Cal/EPA and CUPAs will be making an initial but not comprehensive pass at making these associations. CERS includes various tools for assisting organizations, regulators, and Cal/EPAs in identifying and reassigning facilities to different/new CERS Organizations. However, businesses, regulators, and Cal/EPA will all benefit if as much of this work is done in advance as possible.

Listed below is the suggested approach to this task:

1. Compare the listing of CERS Organizations and their facilities (downloaded in Task 1) with the facilities in the regulator's data system.

FINAL DRAFT

2. For any obvious matches, add the Organization Code to the facility in the regulator's data system. In some cases, while a facility name may share the name of as a CERS Organization (e.g., Chevron, Les Schwab Tires), a specific facility may be a franchise and its reporting is handled by the local owner/operator (rather than the parent organization). Comparisons may also arise where local facility names are ambiguous compared to CERS Organization names. When in doubt, **do not** associate a facility with an Organization. However, regulators should note that facilities not mapped to their multi-facility organizations during the seeding process are likely to just defer the problem to later, resulting in facility transfer requests or organization access requests the regulator and or Cal/EPA will need to process in the future.
3. The CUPA should consider any multi-facility businesses/organizations located entirely within their jurisdiction (e.g., a university, government entity, or regional business with multiple facilities) which do not appear in the downloaded list of CERS Organization Codes. **Well in advance of the regulator's seeding efforts**, the CUPA or PA should encourage these businesses/organizations to request a CERS Organization Code from Cal/EPA by submitting documentation as described in Cal/EPA's October 2011 letter to multi-facility businesses (https://cers.calepa.ca.gov/tempdocs/CERSOrgInfoRequest.pdf).

Step 6: Load CUPA Facility Data into Seeding Template Spreadsheets

Once all existing data in CERS has been reconciled in the regulator's local data system(s) and Organization Codes have been assigned where possible/appropriate, the CUPA's technical staff or contractor should use their professional expertise and familiarity with the regulator local data system(s) to export all of the regulator's reconciled data into the appropriate data seeding template(s). These are:

- Facility Information (equivalent Owner/Operator and Business Activities UPCF)
- Chemical Inventory (equivalent to Chemical Description UPCF).
- Underground Storage Tanks (equivalent to Facility Information and Tank Information forms)

Only the fields listed in these three data seeding templates can be added into CERS via data seeding. Data seeding will not support uploading of any supplemental documentation (e.g., site maps). All CUPA's **must** provide the Facility Information spreadsheet. The Chemical Inventory and UST data seeding templates are **strictly optional** for those CUPAs with sufficiently complete Inventory/UST data.

The CUPA can provide each spreadsheet **only once** for **all** the facilities they wish to seed. Ideally those CUPAs seeding Inventory and UST data would supply their spreadsheets at the same time as their Facility Information spreadsheet. If they don't, as time goes by, the risk increases of new or changed facility information to cause data synchronization problems that will be difficult to identify and potentially time-consuming to fix.

CUPAs should review the checklist at the end of this paper for each seeding spreadsheet they are submitting to Cal/EPA.

Step 7: Send Facility Data Spreadsheet(s) to Cal/EPA

Send the spreadsheet(s) to Chris Allen at callen@calepa.ca.gov along with the name, phone number, and email of who specifically should be contacted when the data is loaded or if there are problems. If the spreadsheets are

FINAL DRAFT

too large to send as email attachments, send information on how to access them on an FTP server reachable by Cal/EPA staff, or contact Chris for further instructions.

Step 8: Spreadsheet(s) Data Loaded by Cal/EPA

Cal/EPA's will review the spreadsheet(s) and ensure proper formatting and then load the Facility Information spreadsheet. Facility records with problems will not be committed to CERS and reported back to the CUPA to resolve problems. Inventory/UST spreadsheets will not be loaded until **all** problems are resolved with the Facility Information spreadsheet.

Step 9: Timely CUPA Follow-up on Data Load as Directed by Cal/EPA

Cal/EPA will contact the CUPA's technical contact with any data load problems and provide a spreadsheet showing the CERS ID assigned to each of the successfully loaded facilities sent by the regulator. Inventory and UST spreadsheets will not be processed until all Facility Information problems are resolved. During this period there may be a certain amount of back-and-forth communication between Cal/EPA and the CUPA's technical contact to resolve problems and negotiate when corrections will be sent and processed. However, Cal/EPA will have very limited staff resources available to perform data seeding, so regulators will be expected to respond and resolve problems in a timely manner or their seeding may be deferred as Cal/EPA continues on with the seeding queue/schedule.

Step 10: Repeat Steps 6-9 for Corrections/Additional Elements as Necessary

Pre-submission Checklist for Data Seeding Spreadsheets

The CUPA should ensure the following issues have been addressed for each seeding spreadsheet before submitting it to Cal/EPA.

- Columns should have the exact column names/headings and order as those of the seeding template.
- All columns/fields shown in the template(s) must be included in the CUPA's submittal, even if they don't provide any data in that field.
- Data types must exactly match those specified in the template/CERS Data Registry.
- All data seeding minimally-required fields are completed for every record (highlighted in green in the templates).
- Any leading and trailing spaces are trimmed from the provided data.
- Do not include any non-standard characters such as tabs, carriage returns, line feeds, etc.
- When possible, provide data in title case (to avoid all UPPERCASE).
- Ensure all dates are properly formatted (YYYY-MM-DD).
- Ensure *FacilityRegulatorKey* values are unique (and will continue to be unique into the future).
- If data in one column repeats from another (e.g., emergency contact is the same as the environmental contact), copy the actual values for each column—**do not** leave blank or show something like "same as environmental contact." Follow this guideline even if it conflicts with instructions on UPCFs.
- Ensure all of the data intended to be seeded into CERS for the CUPA is included in the appropriate spreadsheet. [A few CUPAs may have very large Inventory submissions, in which case multiple Inventory

FINAL DRAFT

spreadsheets can be submitted together. But keep all of single facility's inventory records in the same spreadsheet.]

Questions about Data Seeding

Implementers with questions about data seeding **after** reading this paper can contact Dan Firth (dfirth@calepa.ca.gov) or Chris Allen (callen@calepa.ca.gov).

FINAL DRAFT

Appendix A: CERS Data Seeding Microsoft Excel Templates

Facility Information Seeding Template

http://cers.calepa.ca.gov/Tempdocs/Seeding/FacInfoSeedTemplateNov2011.xls

Use this template to seed data into draft CERS Facility Information Submittal Elements, which consists of the "Business Activities" and Owner/Operator UPCFs.

Hazardous Materials Inventory Seeding Template

http://cers.calepa.ca.gov/Tempdocs/Seeding/InventorySeedTemplateNov2011.xls

Use this template to seed data into draft CERS Hazardous Materials Inventory Submittal Elements, which consists of one or more Chemical Description UPCFs. Facility owner/operators would still need to upload a site map into CERS before seeded Inventory submittal elements could be submitted to the regulator. If a CUPA is seeding a very large number of chemical records making it unrealistic for a single Excel file, the seeding submittal can be sent as more than one file as long as all inventory records for a single facility are in the same Excel file.

UST Seeding Template

http://cers.calepa.ca.gov/Tempdocs/Seeding/USTSeedTemplateNov2011.xls

Use this template to seed data into draft CERS Underground Storage Tank Submittal Elements. The seeding template only captures data from two of the UST UPCF forms: "Operating Permit Application – Facility Information" and "Operating Permit Application – Tank Information." Facility owner/operators would still need to complete several other forms and document uploads in CERS before the seeded UST submittal element could be submitted to the regulator.

FINAL DRAFT

Appendix B: Listing of Fields used in CERS Data Seeding Templates

Appendix B provides an extract from the CERS Data Registry of all fields used in the CERS data seeding templates. Due to its length, it is presented as a separate document at the following web address:

http://cers.calepa.ca.gov/Tempdocs/Seeding/CERS2SeedingNov2011AppB.pdf

Appendix C: Miscellaneous Links to CERS Resources

Regulator Seeding of Facility Data in CERS2 [this document]
November 2011 Draft: http://cers.calepa.ca.gov/Tempdocs/Seeding/CERS2SeedingNov2011.pdf
Final will be published to: http://cers.calepa.ca.gov/Tempdocs/Seeding/CERS2SeedingFinal.pdf

CERS Usage Agreement
See: https://cers.calepa.ca.gov/UserAgreements/ (available sometime in January 2012)

CERS Data Seeding Queue
See: https://cers.calepa.ca.gov/DataSeeding/

CERS Data Seeding web site
https://cers.calepa.ca.gov/DataSeeding/

CERS Data Registry (CDR)
https://cersapps.calepa.ca.gov/DataRegistry/

CERS Electronic Data Transfer (EDT) Information
https://cers.calepa.ca.gov/EDT/

Cal/EPA Letter to Multi-Facility Organizations concerning establishing their CERS Organization
https://cers.calepa.ca.gov/tempdocs/CERSOrgInfoRequest.pdf

FINAL DRAFT

Appendix D: CERS Submittal Elements for Facility Submittals

The table below shows the *Submittal Elements* used in CERS2. Each *Submittal Element* can be submitted (by an organization) and reviewed (by the regulator) independently of each other, except for the *Facility Information Submittal Element*, which must accompany all other submittal elements. To ensure emergency responders have current information and ensure the overall smooth operation of statewide electronic reporting, regulators will need to review all *Facility Information Submittal Elements* within 15 days of being submitted and set their status to "Accepted" or "Not Accepted" (or "Not Applicable" in rare circumstances).

The data that can be supplied by regulators in the CERS Data Seeding Process is highlighted in italics and yellow.

CERS2 Submittal Element	UP Program Element
Facility Information [always managed by the CUPA] • ***Business Activities (UPCF)*** • ***Owner/Operator Identification (UPCF)*** • Locally-Required Document(s)	Business Plan
Hazardous Materials Inventory • ***Inventory: 1+ Chemical Description (UPCF)*** • General Site Map • Annotated Site Map (Official Use Only) • Locally-Required Document(s)	Business Plan
Emergency Response & Training Plans • Emergency Response/Contingency Plan • Employee Training Plan • Locally-Required Document(s)	Business Plan
Underground Storage Tanks • ***UST Operating Permit Application: Facility Information (UPCF)*** • ***UST Operating Permit Application: Tank Information (UPCF)*** • **UST Monitoring Plan (UPCF)** • **UST Certification of Installation/Modification (UPCF)** • UST Monitoring Site Plan • UST Certification of Financial Responsibility • UST Response Plan • UST Owner and UST Operator: Written Agreement • UST Letter from the Chief Financial Officer • Owner Statement of Designated UST Operator Compliance • Locally-Required Document(s)	UST

Onsite Hazardous Waste Treatment Notification • **Onsite Hazardous Waste Treatment Notification: Facility (UPCF)** • **Onsite Hazardous Waste Treatment Notification: Unit (UPCF)** • **Tiered Permitting Unit: Permit By Rule (PBR) (UPCF)** • **Tiered Permitting Unit: Conditionally Authorized (CA) (UPCF)** • **Tiered Permitting Unit: Conditionally Exempt-Specified Waste Streams (CESW) (UPCF)** • **Tiered Permitting Unit: Conditionally Exempt-Small Quantity Treatment (CESQT) (UPCF)** • **Tiered Permitting Unit: Conditionally Exempt - Limited (CEL) (UPCF)** • Tiered Permitting Unit: Prior Enforcement History • Tiered Permitting Unit: Tank and Container Certification • Tiered Permitting Unit: Notification of Local Agency or Agencies • Tiered Permitting Unit: Notification of Property Owner • **Certification of Financial Assurance (UPCF)** • Onsite Hazardous Waste Treatment: Written Estimate of Closure Costs • Financial Assurance Closure Mechanism • Onsite Hazardous Waste Treatment Plot Plan/Map • Locally-Required Document(s)	Hazardous Waste
Recyclable Materials Report • **Recyclable Materials Activities Page 1 (UPCF)** • **Recyclable Materials: Material Page 2 (UPCF)** • Recyclable Materials: Documentation of Known Market • Locally-Required Document(s)	Hazardous Waste
Remote Waste Consolidation Annual Notification • **Remote Waste Consolidation Site Annual Notification (UPCF)** • Locally-Required Document(s)	Hazardous Waste
Hazardous Waste Tank Closure Certification • Hazardous Waste Tank Closure Certification [implemented as document, not a UPCF] • Locally-Required Document(s)	Hazardous Waste
Aboveground Petroleum Storage Tanks • Locally-Required Document(s)	APSA

FINA