Chapter 27: System Management

The registry's instance of the SEER*DMS application and database are hosted at an IMS computer center. All updates to the operating system and SEER*DMS application are managed by the IMS technical team. The database is maintained by IMS database administrators.

IMS system engineers work with local IT staff to implement and maintain a secure network connection from the registry to the SEER*DMS application and database. This connection is typically implemented using a site-to-site or client-to-site VPN.

SEER*DMS uses an HTML-based interface displayed via a web browser. The registry's information technology (IT) staff are responsible for providing on-site technical support for workstations and 3rd party software used with SEER*DMS. A current web browser is the only software required of all users (specifications are provided in *Chapter 3: Using SEER*DMS*). VPN client software may be required in some registries and, in those registries, VPN software may not be required for all staff.

Registry staff may also be required to provide ad hoc programming support. In some registries, this support is provided by research staff and in other registries it is provided by IT staff. The extent to which ad hoc programming support is required varies by registry; the tasks may include writing SQL, writing scripts or programs to manage data files, and generating registry-specific reports. The IMS Technical Support team will provide assistance, but cannot provide large-scale support for these activities.

This chapter provides a broad overview of the duties of registry support staff and a cross-reference to more specific information related to these topics.

In this chapter, you'll learn about

- Technical Skills Required by Registry Support Staff
- Overview of Registry IT Responsibilities
- SEER*DMS Technical Support
- The System Menu in SEER*DMS
- System Tasks
- Patient Set Edits Task
- System Configuration
- System Administration Page
- Reloading the Test Server Database

Technical Skills Required by Registry Support Staff

In addition to proficiency in navigating within the MS Windows environment, the collective skill sets of your staff should include the skills described below.

Ability to Write SQL Queries
Research and/or IT staff must have the ability to write and optimize SQL queries and have an understanding of the SEER*DMS database structure. If the registry has limited SQL experience, formal SQL training should be considered.

SQL should be executed through the SEER*DMS Data Search. In some registries, it may be possible to execute SQL queries using external software (not recommended). The SEER*DMS Data Search provides sample SQL, database diagrams, and field documentation.

Understanding of Data Structures

In order to put the SQL to good use, the staff who write queries must have a full understanding of the SEER*DMS database. For new registries, this would include a mapping to the legacy system, a mapping to the data fields displayed in the SEER*DMS editors, and a complete understanding of all table relationships. Entity-relationship diagrams are available via online help in the SEER*DMS Data Search.

Ability to Create Reports using an External Reporting Package

The SEER*DMS application includes numerous reports that were determined to be useful to all registries and could be implemented generically. If the registry requires additional reports, the registry's IT Staff must be able to generate and maintain the reports in an external reporting package that interfaces with the PostgreSQL database.

Proficiency in an External Programming or Scripting Tool

Registry IT Staff may need to restructure data files into SEER*DMS supported file formats. To do this in a reliable and repeatable fashion, the IT staff may need to write and maintain programs or scripts to perform these tasks. For example, registries often need to write ad hoc programs to create text files in the Generic Supplemental Format containing data sent by non-medical organizations. The files sent by these organizations are typically text files in non-standard formats. Any programming tool may be used for this purpose.

Overview of IT Responsibilities

The following summarizes the responsibilities of the registry's IT staff. Specific information is provided in agreements defined by the registry, IMS, and/or NCI; and in various chapters of this User's Manual, as indicated below.

1. During the transition period, the registry is required to provide data to IMS for migration into SEER*DMS. These data need to be in well-defined file layouts; and the files must first be encrypted and then transferred using secure protocols.

2. Using the SEER*DMS system itself, registry IT staff will:

    a. Extract data from the SEER*DMS database and save the data in the format required for a data submission, ad hoc data requests, and extract files for use in analytic tools. Standard extracts for SEER, NAACCR, NPCR, and other submissions are available as system extracts. Registry-defined extracts can be created through the SEER*DMS interface; templates for registry-defined extracts are also available (see *Chapter 24: Creating Reports and Extracting Data*).

    b. Import data files as described in *Chapter 5: Importing Data Files*.

    c. Provide technical support to registry staff, as needed. This requires a general knowledge of the system features described throughout this manual.

3. Staff may need to create custom formatted reports in some registries. This will require:

    a. A working knowledge of SQL and external reporting software (e.g., Crystal Reports).

    b. Acquire and install the reporting software on the PCs of all registry staff who need to run external reports.

4. Using Windows tools, registry IT staff will:

    a. Configure the registry's personal computers and maintain current versions of client software required for SEER*DMS as described in *Chapter 3: Using SEER*DMS*.

    b. Provide on-site technical support for all PC and network-related issues.

    c. Write, test, and execute routines to pre-process data files that are not formatted per registry-defined specifications for SEER*DMS import files.

5. Write, test, and execute routines to process extracted data files (running edits, making data files available to staff, submitting data to NCI and other organizations).

SEER*DMS Technical Support

Critical issues that inhibit registry operations should be reported immediately to IMS staff. Non-critical issues should first be triaged by the registry's local technical support team.

The registry should appoint a local technical support team who can respond to routine issues. This team should include representatives from IT, editing, and management. The registry team will triage routine technical support issues and perform an initial investigation. If an issue cannot be resolved on site or requires a system change, provide a description of the problem and the results of the investigation to IMS. Issues and requests for new features should be submitted via the Technical Support Squish project (https://www.squishlist.com/seerdms/support).

The System Menu in SEER*DMS

The System menu provides access to several modules related to SEER*DMS system management. The list below shows the items in the System menu, a brief description, and the chapter of the SEER*DMS User's Manual with more detailed documentation.

- Roles – System permissions control the user's access to system functions and data. A role is a set of permissions; roles can be assigned to user accounts. Registry staff create and maintain system roles. For more information, see *Chapter 26: System Roles & Permissions.*

- User Accounts – User accounts provide login authentication, access controls, and the ability to track user activity. Registry staff create and maintain user accounts. For more information, see *Chapter 25: User Accounts.*

- Edits Manager – Edits are scripts to validate data fields. Edits from several standard setters are available in SEER*DMS, these include edits from SEER, NAACCR, NCDB, NPCR, etc. Registry staff can use the Edits Manager to turn edits on or off; set conditions for when edits run; or write registry-specific edits. See *Chapter 7: Edit Errors.*

- Imports – Use the Import Manager to load data files into SEER*DMS, to monitor the results of imports, and to create records via Data Entry. See *Chapter 5: Importing Data Files* and *Chapter 6: Data Entry.*

- Keywords – A keyword is a significant word or phrase that may appear in text data fields. SEER*DMS includes a list of keywords typically found in medical records related to cancer diagnosis and treatment. SEER*DMS uses color-coded highlighting to emphasize the keywords when the text fields are displayed on the screen. The System > Keywords menu item allows registry staff to add and delete words and phrases from the keyword list.

- Reference Data – Reference data sets may be defined by registry IT staff and may include patient data stored in the SEER*DMS database or data from external data sources. Reference data that match a patient are displayed in the Patient Set editor (there is a tab in the right panel of the Patient Set editor that is only shown when reference data are defined). Custom matching algorithms may be defined by the registry. If your registry is interested in using Reference Data, please contact SEER*DMS Technical Support.

- Workflow – Workflow routing and assignment rules can be maintained in the Workflow Task Manager. The routing scripts control the processing of records as they exit the automatic match task. Values of data items and the results of the matching algorithm can be used in the scripts. Using these scripts, the workflow can be modified at any time without requiring a new release of the SEER*DMS application. The workflow manager also allows the

registry to automatically set worklist flags. Flags provide a way to identify problem tasks and to set a priority to tasks.   For more information, see *Chapter 4: Using the Worklist*.

- Tasks – The system tasks are batch processes executed against the database. These are described below, in the *System Tasks* section of this chapter.
- Configuration – The configuration manager lists parameters which define registry-specific settings. The configuration manager provides a mechanism for changing system settings without requiring a release of the SEER*DMS application. Registry staff with the *system_administration* permission have read only access to the configuration manager. Documentation for each parameter can be viewed by clicking the parameter.   At this time, configuration settings can only be changed by IMS staff.
- Administration – The System Administration page provides access to system logs, system files (XML files used by various SEER*DMS modules), and tools used by the SEER*DMS development team to monitor system performance and investigate technical support issues. For more information, see the *System Administration* section of this chapter.

System Tasks

The Tasks section of the SEER*DMS System menu provides access to utilities or "tasks" to execute processes in batch mode. For example, the Patient Set to Patient Set Matching task applies the duplicate patient matching algorithm to the full database. The algorithms are applied each time a patient set is created or saved, therefore, it is only appropriate to run this system task when the algorithms are changed.

The most commonly used system task is the Patient Set Edits task. Edits are executed each time a patient set is opened or saved, therefore, it is only necessary to execute the edits against the full database when the edits are updated. The Patient Set Edits task also provides a mechanism for executing individual polishers on the patient sets. This is typically done when a new polisher is added or changes are made to a polisher. More information and step-by-step instructions for the Patient Set Edits Task are provided in the *Patient Set Edits Task* section of this chapter.

The system tasks available in SEER*DMS vary by registry and other requirements. For example, a system task may be added to implement a conversion routine related to a new version of Collaborative Stage or new NAACCR requirements. These tasks are then removed when they are no longer needed. Documentation for a task can be viewed by selecting System > Tasks and clicking on the name of a task.

A task can only be executed by a user with the appropriate system permission; and some tasks are restricted to IMS staff. For more information related to a specific task, please submit a request to the SEER*DMS Technical Support team.

Patient Set Edits Task

All edits are executed each time a patient set is opened, validated, or saved in the SEER*DMS editor. The Patient Set Edits system task enables you to re-execute the edits on patient sets in the database. You may run the edits on all patient sets, on a cohort defined by year of diagnosis, or on a list of Patient Sets. Use the Patient Set Edits task to ensure that new or modified edits are evaluated. Note: It is not necessary to run the task when an edit is deleted or deactivated. The edit errors for a deactivated/deleted edit are removed immediately when the edit is saved.

A polisher is a system utility that derives, calculates, or assigns data field values. Polishers are used to derive collaborative stage variables; assign census tract based on address, calculate the age at diagnosis based on date of birth and date of diagnosis, etc. When a patient set is opened, saved, or validated, a polisher will be executed if the value of a related data item has changed. You have the option of executing up to six polishers in the Patient Set Edits task. Use this feature

with care. Polishers change data and some polishers are not appropriate for data migrated from a legacy system. A polisher should not be executed across the database without careful consideration.   Please consult with the SEER*DMS technical support staff and review the results of the task on the test server before executing a polisher in production.

If one CTC meets the criteria defined by the parameters, the edits are executed for every CTC within the Patient Set. If a polisher for a CTC-level field is executed then that polisher will only be run on CTCs with a year of diagnosis that meets the criteria defined by the parameters.

To execute the Patient Set Edits task:

1. Click **System > Tasks.**

2. Click the **Patient Set Edits** link.

3. Define the Patient Sets considered by the task:

    a. To run the task for all non-deleted Patient Sets: do not enter a **Start** or **End Year**; and set **Include Unknown Year** of diagnosis to Yes. To limit the task by year of diagnosis:

        i. Define a year range using the **End Year** and **Start Year** parameters. You may define it as an open-ended range.

        ii. To include data with unknown year of diagnosis, set **Include Unknown Year** to Yes.

    b. If you prefer to specify a set of Patient Sets, paste a list Patient Set IDs into the **IDs** box. The list can be space or comma separated; the PAT- prefix is optional.

4. To execute a polisher, select a polisher from one of the **Extra Polisher** drop-down lists. Warning: consult with IMS technical support or review the results of executing the polisher on the test server before executing a polisher on the production database.

5. You may enter text related to this task in the **Comment** field. The comment for the last execution of the task is stored in the database (utility_history table).

6. Click **Start**.

The edits will be re-evaluated for each patient set in the cohort. Deleted patient sets are completely ignored by the task, that is, edits are not evaluated and polishers are not executed on deleted patient sets.

In order to avoid creating an inordinate number of worklist tasks, a Resolve Patient Set Errors task will *not* be created for each patient set with an edit error. If the logic of a new or modified edit is implemented incorrectly, this could erroneously create tasks for a large number of patient sets. However, existing Resolve Patient Set Errors (RPSE) tasks are re-evaluated. If no edits are failing for a Patient Set in an RPSE task, the task will be closed.

You may use the Edits dashboard, the Data Search, or reports to identify patient sets with errors and to evaluate the error levels in the patient set data. The Data Search can also be used to create Resolve Patient Set Errors tasks for Patient Sets that are failing a new or modified edit.

Two system reports are available for identifying the edit errors that were triggered and the patient sets that are involved:

- *RPT-064A: Frequency of Edit Errors in the Patient Set Data-* Run this report to evaluate the error levels in the patient data. In order to verify that modifications to edits, polishers or data did not have unexpected results, set the parameters to generate frequencies of errors for all edits in all patient set data.

- *RPT-064B: Patient Sets with Edit Errors* - Run this report to obtain a listing of Patient Sets with an error related to a particular edit.

Instructions for creating reports are provided in *Chapter 24: Creating Reports and Extracting Data*.

Aborting a System Task

Requires system permission: *worklist_task_terminate_auto*

You may use the worklist to stop the execution of a system task. However, all changes made by the task may not be reversed. Some system tasks that update data perform the updates in batches. Updates that were made before you click Abort Task will not be reversed. You may use reports or query audit logs, the worklist, and other data to determine what was changed.

To abort a System Task:

1. For system tasks, the automatic task will be assigned to you and a link will be shown in the My Tasks section of your home page. Click the Automated task link for the appropriate task type to access the worklist.

2. Click the **Task ID**.

3. Click **Abort Task**.

4. Click **OK** to confirm.

System Administration Page

Requires system permission: *system_administration*

The System Administration page is a set of tools that allows the IMS technical team to monitor system performance and investigate technical support issues. Registry staff use the information in the server log to monitor the status of system tasks or large imports.

To access the tools listed below, select **System > Administration**. When you first enter the page, the Module drop-down will automatically be set to Logs and the current server log will be displayed. Each module is described below.

Benchmarks
A tool to evaluate system performance by executing standard processes such as loading and initializing patient sets. This tool is one of many tools used by the IMS technical team to evaluate program changes and to investigate issues related to system performance.

Database
This module is a diagnostic tool that allows the SEER*DMS development team to monitor database efficiency. Queries are executed which generate statistics. These statistics are useful in monitoring the state of database and identifying inefficient queries.

Environment
This module lists environment variables in the operating system.

Login History
The username, login date, logout date, monitor resolution, IP address, and User Agent are listed for SEER*DMS sessions. The logout date may not accurately reflect user activity. Many users close the web browser without logging out of SEER*DMS; their session will remain open until the period of inactivity exceeds the *system.session.timeout.seconds configuration* parameter. The default setting for the timeout parameter is two hours.

Review the text in the User Agent column to verify that users in your registry are using current versions of their web browsers. SEERD*DMS users should enable the browser's auto-update option, if permitted by registry and local policies that govern the maintenance of software on registry PCs.

Logs

The Logs module provides access to server logs and user access logs. The server logs include Info, Warning, and Error messages from the SEER*DMS application and JBoss server. IMS staff continually monitor the server logs and are automatically notified when an error is written to the log.

The "extra" server logs contain detailed information written by the SEER*DMS application. Typically, the entries written to the extra logs are related to the processing of records in the workflow. This information is used by the IMS technical team to evaluate workflow processes and investigate workflow issues reported by registry staff.

The access logs provide a detailed history of user activity (IP address, username, timestamp, and the URL of the SEER*DMS page that was accessed). An entry is added to the log each time a user changes pages within SEER*DMS. The primary purpose of the access logs is to provide a way for the IMS technical team to identify the steps a user had taken when they encountered a system error. These logs are also used by SEER*DMS reports related to staff productivity. To view these reports, go to View > Reports and set the Category to Productivity.

Use the Logs drop-down menu to search for a log based on the log date and type listed in the filename (server_log, access_log, or extra-server log). This drop-down menu also provides options for searching All Logs, All Access Logs, All Server Logs, or All Extra Logs. The search engine returns the most recent log entries that match the search. The number of results returned is limited by the # Rows drop-down.

The maximum number of lines displayed in the viewer is based on the value in the # Rows drop-down. If a file has more than the maximum number of lines then the last segment of the file is shown. To review early entries in a log file that exceeds 5,000 rows, download the file to your PC and open the file in a text editor.

Memory

The Memory module shows current memory allocation, memory usage, and memory pool status indicators.

Response Data

The Response Data module provides metrics related to SEER*DMS response time, that is, the elapsed time between the end of a user request and the display of results in SEER*DMS. The SEER*DMS technical team monitors the response time for all aspects of the system including the loading of pages in the Patient Set and Record editors, the display of results in the Data Search or Patient Lookup, etc.

Sessions

The Sessions module shows the username, overall length of time, length of time since the last activity, IP address, and User Agent for each active session. A session is active until the user logs out or the period of inactivity exceeds the *system.session.timeout.seconds configuration* parameter.

Statistics

The statistics module summarizes the execution time of various activities (execution of edits and automated workflow tasks, for example). This module is used by SEER*DMS technical support staff to investigate issues related to system responsiveness. The workflow section shows statistics for automated tasks based on the workflow history table. Entries for automated tasks are purged from the workflow history after 30 days.

System Files

The system files module provides access to XML files used by the SEER*DMS application. A data-driven programming approach is used throughout SEER*DMS. The XML files listed in the System Files module support the following modules: Edits, Hibernate, Importer, Lookups, Matching, Reports, Screening, and the Workflow.

The system files are used by the IMS team to review registry-specific logic. Typically, these files are not useful for registry staff. However, the SQL for every SQL-based system report is available in the Reports section of the system files. This SQL can be used as a template if you are writing a report that is similar to one of the system reports. Note: SQL is not available for Java-based reports.

System Properties

The system properties are the Java JDK settings.

Test Server

The Test Server reload script can be triggered from this page. This is only available in the production application. For more information, see the *Reloading the Test Server Database* section of this chapter.

Threads

This viewer shows the concurrent system processes that are executing or waiting.

Workflow History

This workflow history module is a log of completed workflow tasks. Automated tasks are shown in gray and manual tasks are shown in black.

Reloading the Test Server Database

Requires system permission: *test_server_reload*

Each registry has two versions of SEER*DMS, one is for production work and the other is available to test new registry procedures, to test new linkages that might have a significant impact on the registry's workflow, and to test beta-release versions of SEER*DMS. Access to the SEER*DMS test server is restricted to users with the *test_server_access* system permission.

The database for the SEER*DMS test server is completely separate from the production database. The test database can be reloaded by copying the current production database.

To reload the database for the registry test server:

1. Login to the production version of SEER*DMS.

2. Select **System > Test Server**.

3. Click **Schedule Reload**.

A small file indicating that the request was submitted will be created on the server. A Cron job on the server runs periodically and checks for these requests. An email will be sent to you indicating that the reload has been scheduled and a second email will be sent when the reload is complete. The database is typically reloaded within 30 to 60 minutes. The login page may not be available until a few minutes after you receive the email indicating that the database has been reloaded. The application server is restarted after the reload and the restart may take 5 to 10 minutes.

This is only available if SEER*DMS is hosted at an IMS data center. Registries with local servers must continue to login to the server to run the reload script. Note: All SEER*DMS registries are expected to be hosted on IMS data centers by the end of 2015.