# IRIS Document Migration

## Background

The documents associated with R2D2 consent applications and compliance inspections are currently stored in the Herman document repository in [\\file\herman\E\RM\04\00](file:///\\file\herman\E\RM\04\00), with a second repository created after the initial IRIS go-live at [\\file\herman\E\RM\08\02\Consent Monitoring IRIS](file:///\\file\herman\E\RM\08\02\Consent%20Monitoring%20IRIS). These files need to be migrated into the document repository associated with IRIS Authorisations or Regimes.

In the R2D2 system, files are stored in a folder named as per the consent authorisation number including categorised sub-folders. The newer Compliance Monitoring folder name is the authorisation business identifier.

In IRIS, files save to a folder named as per the business identifier for the application, authorisation or regime. In the case of contacts and other IRIS objects that do not have a business identifier, the folder is named as per the IRIS ID for the contact (as displayed in the URL when viewing the contact details).

There are no categorised sub-folders in our IRIS document system. The document migration process needs to add metadata to the filename, generally as a prefix. So the name of the copied file will not exactly match the source file. The migration process is a *copy, not a move*.

## Process

### Configuration

The migration requires an initial configuration in which four parameters are set:

#### IRIS Database Server

The name of the server where the IRIS database resides. This will be one of:

* **dbserver** for the final document migration
* **iristestdb** or **irisdevdb** for any test migrations

Enter a value into the **SERVER** textbox and press **Enter** to activate a validation process. This will populate both **DATABASE** dropdown lists with the collection of database names on the identified server and pre-select the IRIS database and mwrcdb database (disabled fields as these values are hard-wired to the server name).

Click **Step 1: Create Workspace** in order to prepare the database tables required for holding the following migration parameters. Textboxes and buttons for entering the folder names become disabled until this task is complete.

#### Migration Identifiers

This is a combination of numeric and textual identifiers for the particular migration taking place. It should be set in the Document Migration Process table using SQL (all things going to plan, rows will be created for this during the regime migration process that runs in SQL Management Studio.

The Migration ID is an integer representing the main migration event that is occurring and must match the migration ID used for the main event that creates the objects and relationships in IRIS. The File Migration ID refers to a particular source-destination combination for the files to be copied. The File Migration ID must be unique within the column.

#### Source Documents Root Folder

The path to the folder where the source documents are located. For the final migration, this will probably be [**\\file\herman\E\RM\04\00**](file:///\\file\herman\E\RM\04\00) but for testing it can be any folder where a sample set of R2D2 consent documents are located. There is also a source repository of documents saved subsequent to the initial IRIS Go-Live, [\\file\herman\E\RM\08\02\Consent Monitoring IRIS](file:///\\file\herman\E\RM\08\02\Consent%20Monitoring%20IRIS).

Each folder has a checkbox setting to indicate if the root-level folder names refer to R2D2 legacy ID values, or IRIS business.

#### Destination Documents Root Folder

The path to the folder where the documents will be migrated. For the final migration this will be [**\\file\herman\D\IS\01\01\IRISLIVE**](file:///\\file\herman\D\IS\01\01\IRISLIVE) but for testing can be any convenient location.

### Migration

The actual migration process is a four-step operation. For each configured File Migration.

#### Step 1: Create Workspace

Click this button to populate the required tables in the **Workspace Database**. There are four workspace tables that have been created on the **hrc** schema in the **IRIS\_Migration** database on the nominated server.

|  |  |
| --- | --- |
| * DocumentMigrationProcess | One record naming a migration ID, source and destination folders |
| * DocumentMigrationProcessTask | A logging table indicating which steps have been completed |
| * DocumentMigrationFolders | A recursive list of sub-folders identified for migration within the source folder populated in step 2 below |
| * DocumentMigrationFiles | A list of files included in the migration with a result status and (possible) error message, populated on completion of step 3 below |

#### Step 2: Compile Mapping Data

This step will recursively scan all the sub-folders of the nominated source directory and populate the DocumentMigrationFolders and DocumentMigrationFiles tables with the names of folders and files to be included in the migration process.

This step will take several minutes to complete on the final run because it requires scanning over 116,322 folders and applying a set of business rules to select folders whose contents will be copied to the IRIS folders. These folders contain 275,311 files totalling around 292GB of storage.

#### Step 3: Scan for Duplicate Files

This step scans all the files associated with each regime, comparing file name and size. The process tags matching files as potential duplicates. The process subsequently physically loads each potential duplicate and calculates a CRC (Cyclic Redundancy Check). Any files that have the same name and size, but differ in their CRC are copied and versioned, else only one of the copies will be transferred.

#### Step 4: Copy to New Location

This step will use the mapping tables to complete copying contents of all identified folders from the source directory into the destination folders. On a test run of all documents this took around four hours to complete. Note that the process *does not copy all files*. The mapping process identifies files associated with applications and authorisations that meet a set of criteria depending on the folder name.

As the process proceeds permissions on files in the *source* folder, and the folder itself, are reset to read-only for everyone. Permissions on the new copy are set to full control for everyone.

The application also records the results of the copy for each file and updates the DocumentMigrationFiles table during the process.