XML Documentation for Adobe Experience Manager Content Migration Guide

Adobe

December 10, 2019

Contents

About this guide

XML Documentation for Adobe Experience Manager (referred to as *XML Documentation solution* later in this guide) is a powerful, enterprise-grade component content management solution (CCMS). It enables native DITA support in Adobe Experience Manager, empowering AEM to handle DITA-based content creation and delivery. It empowers authors to create content using any offline DITA authoring tool, such as Adobe FrameMaker or an easy-to-use built-in Web Editor.

This guide provides the best practices and other relevant information about how you can migrate your DITA and non-DITA content to Adobe Experience Manager and then use XML Documentation solution for authoring and managing your content.

Migrating DITA content

If you are migrating your DITA content, follow the guidelines and steps mention in this document.

Migrating non-DITA content

If you have non-DITA content, you need to convert the content to DITA by following the steps mentioned in the "Migrate non-DITA content" section in the document XML Documentation for Adobe Experience Manager Installation and Configuration Guide, and then you can follow the guidelines and steps mention in this document.

Migration process

The migration process can be broken down into the following three stages:

*   Pre-migration stage: this stage covers the preparations required for the content migration.
*   Migration tasks: this stage covers the content upload and configurations.
*   Post-migration verification: this stage covers the verification required after the content upload.

Pre-migration phase

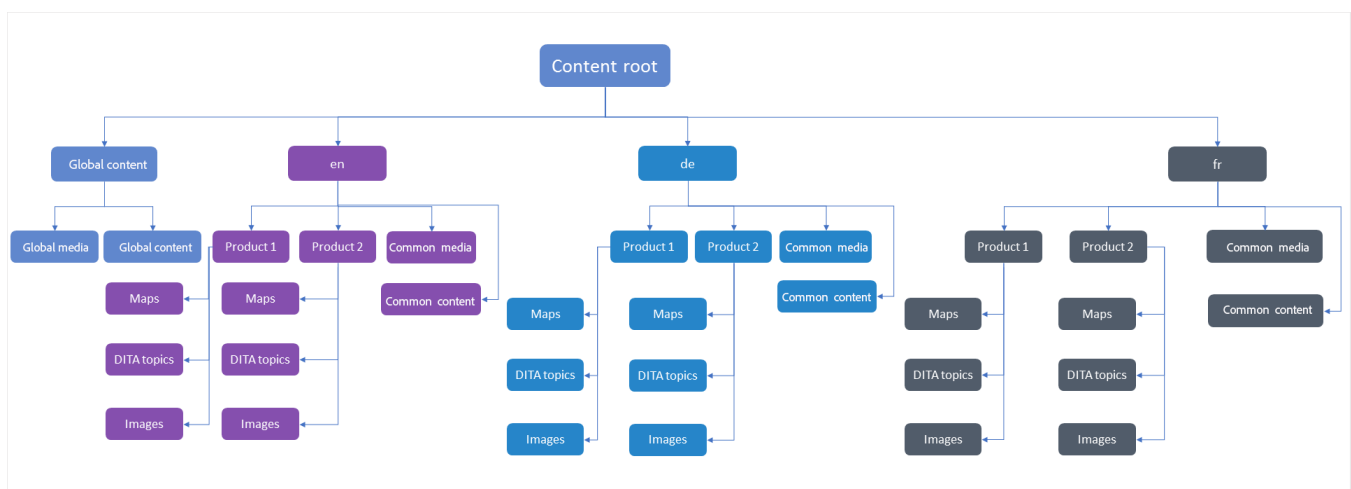The following sections cover the required pre-migration tasks.

Create folder structure in AEM

If your current content management system organizes the content in a flat folder structure, you need to create and organize the content in a folder structure in AEM.

Consider the following point while creating a folder structure in AEM:

- The folder name must comply with the file naming standards as described in *Comply with the file naming standards* section.

- If you translate content in different languages, you must create folders corresponding to each language. Each of these language folders will contain the content corresponding to that language. For example, you can create folders using the language designator like `de` for German, `fr` for French, and so on. Or, you can create folders using the language and region designators like `fr-FR` for French as used in France or `fr-CA` for French as used in Canada.

- The globally used content and media that don't require localization, should be kept out of the language copies.

- All the common content that must be localized should be kept in a common folder under the language folder.

- A folder should not have more than 1000 files in it.

The following diagram shows an example of a folder structure in AEM when you have globally used content and three language copies.

Prepare your content for migration

Clean CCMS-specific markup

Remove all the CCMS-specific markups from your content that does not comply with DITA standards. For example, use actual image file references if your content references images with `.image` links.

Comply with the file naming standards

Make sure that the filenames are according to the following guidelines:
- There should be no spaces, apostrophe, braces, or equals sign in filenames to simplify file references in your content.
- All references in your files should be valid URLs. Having spaces in a filename will require references to be URL-encoded.
- Ensure that there are no special or non-ASCII characters in filenames. Your filename should only contain alphanumeric characters with underscore and hyphens.
- Ensure that the DITA map files are saved with `.ditamap` extension.

Identifying images to be uploaded

XML Documentation solution supports multiple source image formats like `.ai` and `.psd` (Adobe Illustrator and Photoshop) in addition to web-enabled image formats like `.eps`, `.svg`, `.jpeg`, `.png`, `.tif`, and so on.

Also, the solution supports creating different sizes and resolutions of images from the source content dynamically. Therefore, you don't have to upload the different resolutions of images manually into the repository. The system will automatically create the configured renditions every time you upload a new image or update an existing one.

If you decide to upload the source image formats (instead of the different resolutions), you need to update the image references in your content to point to the source format rather than their renditions.

Handle globally reused content and media

Things to consider for globally reused content and media files:
- Identify all the globally referenced content (for example, variable, snippets, images, or videos) that is used in your documentation.
- Decide where do you want to store it in the AEM directory. For example, if you don't have a need to localize the content, it should be stored outside a language folder. If you want to localize it, then store it inside a language folder.

Handle snapshots

Consider the following points when migrating versioned content:

• Migrating the last released version(s) of your documentation (in addition to the latest version) can be easily achieved in XML Documentation solution. To migrate versioned content, you need to upload the content corresponding to each version individually into the AEM repository.

• Each time you upload your content in AEM repository, it is automatically versioned in AEM. This allows you to access and publish any older version of your documentation, as and when needed.

• To create a clean versioned content history, you should migrate the different versions in chrono-logical order.

• When migrating a documentation version to AEM, you need to upload all DITA content and its dependencies (including maps, topics, and media assets) corresponding to that version.

Handle metadata to be migrated (if any)

If you want to migrate any file-level metadata, map the fields of the current metadata to the fields of the new system. Use the existing metadata fields or create new one which maps to your metadata.

For example, you may want to migrate the details of the owner or author of a topic to the new system to keep a track of the content owner or author.

NOTE: You can use the Curl or the Rest APIs to add metadata. For more information about the Curl APIs, see "Use curl commands" section in the document XML Documentation for Adobe Experience Manager Installation and Configuration Guide. For more information about the Rest APIs, see the https://helpx.adobe.com/experience-manager/6-3/assets/using/mac-api-assets.html.

Export DITA-OT and DTDs

If you are using any custom DITA-OT or specialized DTD, you need to export the DITA-OT and the DTDs from your system so that you can later configure them in the new system.

Migrate content and configurations on a test server

The following sections cover the migration tasks that you need to perform on a test server.

Configure custom DITA-OT and DITA specialization

If you are using any custom DITA-OT or DITA specialization, configure that in the new system. To know about how to configure the DITA-OT and DITA specialization see the "Use custom DITA-OT and DITA specialization" section in the document XML Documentation for Adobe Experience Manager Installation and Configuration Guide.

Configure document states

If you are using custom document states to signify at what state the document is placed in the document life cycle, then such document states must be configured in the new system. Configuring document states will ensure that the content is assigned the correct document state once it is uploaded in AEM repository. To know about how to configure document states see the "Configure document states" section in the document XML Documentation for Adobe Experience Manager Installation and Configuration Guide.

Configure image rendition for output generation

AEM comes with a set of default workflows and media handles to process assets. In AEM, there are pre-defined workflows to handle asset processing for the most common MIME types. Typically, for every image that you upload, AEM creates multiple renditions of the same in binary format. These renditions may be of different size, with a different resolution, with an added watermark, or some other changed characteristic. For more information about how AEM handles assets, see Processing Assets Using Media Handlers and Workflows in AEM 6.3, 6.2, and 6.1.

Configure default output presets

If you have configured output presets for generating output, then those must be configured in the new system. Output presets can be configured at global level or folder-specific profiles can be created to store output configurations. To know about how to configure default output presets see the "Configure output presets" section in the document XML Documentation for Adobe Experience Manager Installation and Configuration Guide.

Upload content

The recommended approach for uploading content is to upload content into a working directory in AEM. It must be in the same directory structure as it exists in your current system. Make sure that you upload content in batches of maximum 500 topics in each batch with a gap of 5 minutes in-between two batches. This enables the post-processing workflow to complete and prevents the upload process from getting stalled.

If you are working with more than 500 topics, then it is recommended to use a tool that allows you to specify a batch size and a throttle between batches. Once such tool is the CSV Asset Importer tool. This tool allows you to define a batch size for each upload task along with time gaps between two tasks. To get more details about the tool, see the CSV Asset Importer tool page.

During the content upload process, ensure that there are no other CPU-intensive processes running. It is recommended to allow the upload process to complete before starting any other task on the AEM server.

After uploading your content, re-organize content using the AEM Web UI (as described in *Create folder structure in AEM*). The XML Documentation solution will automatically fix all the references as you move content around.

Alternatively, you may create scripts to update the references before you upload the content into AEM directly into the new directory structure. This reduces the manual work if you have to upload multiple snapshots/languages.

NOTE: You can use a WebDAV tool like WinSCP to upload content less than 500 topics. Or, you can also use the Curl or the Rest APIs for the same. For more information about the Curl APIs, see "Use curl commands" section in the document XML Documentation for Adobe Experience Manager Installation and Configuration Guide. For more information about the Rest APIs, see the https://helpx.adobe.com/experience-manager/6-3/assets/using/mac-api-assets.html.

Upload a snapshot

Generally, you would have content with multiple versions or snapshots. For example, there could be a snapshot of the very first version (1.0) of a guide stored in your file system. Now, there could be a latest version (5.0) available for the same guide. If you want to take such content and maintain their version history, then it is recommended to start uploading the content starting from the very first version of the content going up to the latest version.

Perform the following steps to upload a snapshot of your content:

1) If your content includes images, then you must disable the DAM Update Asset workflow before starting the upload process.

2) Upload all content (topics, maps, images, and so on) corresponding to a snapshot in a working directory and reorganize content in the new directory structure.

3) *(Optional)* Create a baseline for the content to be able to publish it at any later point.

   For more information about using the baseline feature, see "Use Baseline for publishing" section in the XML Documentation for Adobe Experience Manager User Guide.

4) Add labels to each snapshot corresponding to the release. For example, if you have created a content snapshot for release v1.0, you can add v1.0 as a label to the content.

For more information about using the label feature, see "Use Label" section in the XML Documentation for Adobe Experience Manager User Guide.

5)   If you need to upload a snapshot of the localized content, upload the localized version of the snapshot and repeat all the above steps.

NOTE: Make sure that the localized content follows the same directory structure (as defined in source content) in all language copies. This ensures that the language copies are in sync with the source copy. Also, ensure that the file names are same as in the source content so that the system can relate them as language copies.

6)   Verify the validity of uploaded content. See *Verify content* for more details about content verification.

Upload the latest content

After uploading all snapshots of your content, the next step is to upload the latest version of the content. Again, you should use a tool such as the CSV Asset Importer for topics more than 500 or WinSCP for less than 500 topics. Ensure that you are uploading content in batches and there's sufficient time gap in between two batches to allow the active batch to complete.

If you have localized content, then follow the same procedure as described in Upload a snapshot and upload all language copies of your content.

NOTE: Make sure that the localized content follows the same directory structure in all the language copies. This ensures that the language copies are in sync with the source copy.
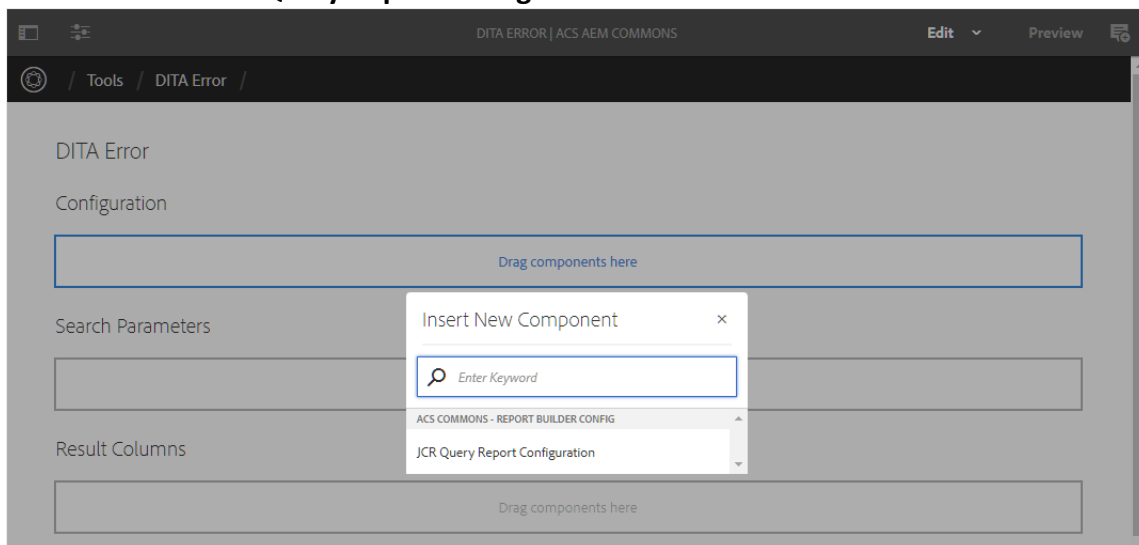
After uploading all required content, verify the validity of uploaded content. See *Verify content* for more details about content verification.

Verify content

You can validate your content using any of the following methods:

•   Create, open, and edit the DITA topics.

For more information about it, see sections "Work with the Web editor" and "Work with the Map editor" in the XML Documentation User Guide.

•   Check for broken links using the Report tab of the DITA map console.

For more information about it, see the section "DITA Map Report" in the XML Documentation User Guide.

•   Generate output.

For more information about it, see sections "Output generation" in the XML Documentation User Guide.

•   Use ACS Commons to generate a content validation report.

Perform the following steps to generate the report:

a)   Download the ACS Commons package from here.
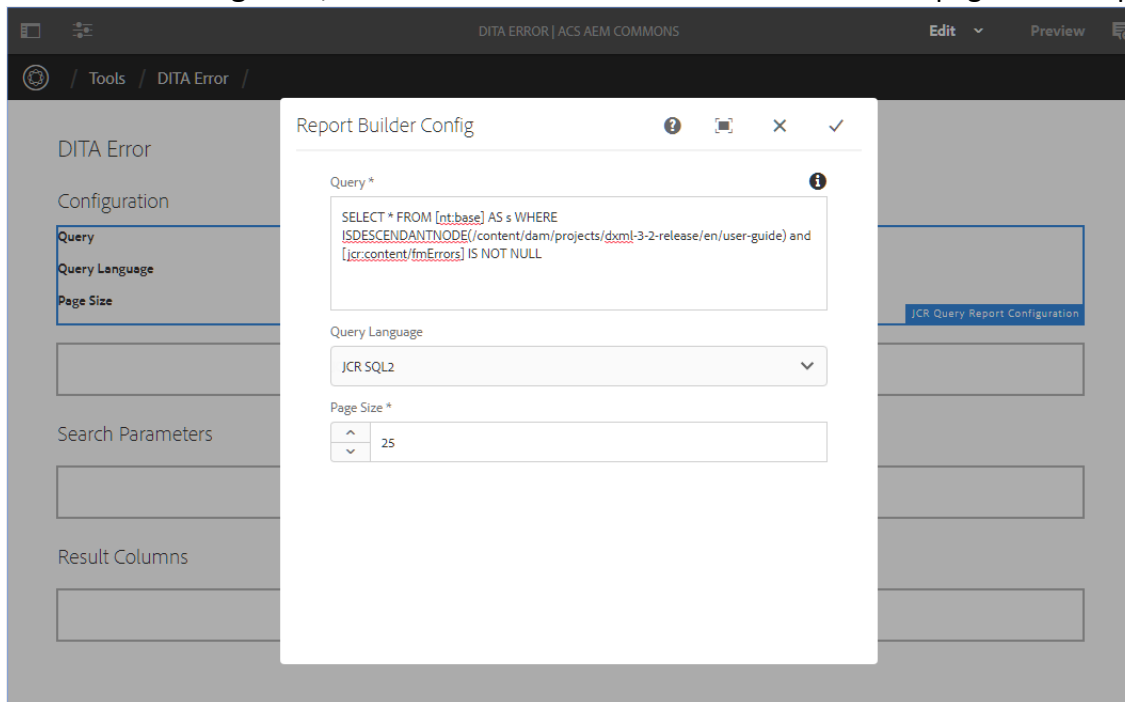
b)   Install the ACS Commons package on the AEM server.

c) Click on the Adobe Experience Manager link at the top and choose **Tools**.

d) Select **ACS AEM Commons** from the list of tools.

e) Click on the **Reports** tile.

f) Configure the Reports page as:

- Enter a title for the report and click **Add Report**.

- In the *Configuration* section, click **Drag Components Here** > **Insert Component**.

- Select **JCR Query Report Configuration**.

- Click on the new JCR Query Report Configuration component and select **Configure** from the quick action menu.

- In the *Query* field, enter the following query:

```
SELECT * FROM [nt:base] AS s WHERE
ISDESCENDANTNODE([add-your-content-path-here]) and
[jcr:content/fmErrors] IS NOT NULL
```

- In the **Page Size**, enter the number of records to show on each page of the report.

- Click the done checkmark.
- In the **Result Columns** section, add two columns with the following details:

Column type: ACS Commons Report Builder Path Column	**Heading**: Path	Select the **Include Link** option
Column type: ACS Commons Report Builder Text Column	**Heading**: Error	**Property**: `jcr:content/fmErrors`

- Click on the report title at the top.
- Run the report by clicking either **Execute Report** or **Download Report** button.

The report contains a list of all topics that have errors.

Create a package of content and configurations

Once you have uploaded your content and configured XML Documentation solution on a test server, you need to take a backup of the content and configurations. A backup of content and configuration can then be easily deployed on your production server.

If your content is more than 500 topics, then it is recommended to use the VLT Tool to move your content from the test server onto the production server. For content lesser than 500 topics, you can create a package of your content as explained in the following procedure.

Perform the following steps to package your content and configurations:

1) Open CRX Package Manager and click **Create Package**.

2) Enter a name for the package. For example, if you are creating a package for your content, name the package as `content-product-user-guide`.

3) *(Optional)* Enter a version number for the package. For example, if you are creating a package for version 1.0 of your content, then enter `v1.0` in the **Version** number.

4) Use the default **Group** of `my_packages` and click **OK**.

   The package is added and shown at the top of the list of packages.

5) On the package name, click **Edit**.

6) On the **Edit Package** dialog, select **Filters** and click **Add filter**.

7) Select the node where your content is stored in the AEM repository and click **Done**. Typically, the content is stored in the `/content/dam/` node.

8) Click **Save** to save your changes.

9) Click **Build** to create the package.

10) Perform Steps 2 to 9 to create a package for your configurations. XML Documentation solution saves configurations at different locations. The following table list the locations where you can find various configurations:

Configuration for	Location
DITA profiles	`/content/fmdita/profiles`
Folder profiles	`/content/fmdita/folderprofiles`
Conditional profile	`/content/fmdita/conditionalprofiles`
Default output presets	`/content/fmdita/config/presets.json`[1]
Map collections	`/content/fmdita/mapcollections`
Document states	`/content/fmdita/states`
Configuration Manager	`/apps/system/config/com.adobe.fmdita.config.ConfigManager.config`
XML Editor	`/apps/system/config/com.adobe.fmdita.xmleditor.config.XmlEditorConfig.config`

1. `presets.json` file is available on versions 3.0 or earlier. If you are using XML Documentation solution version 3.1 or later, you can ignore this file

11) After your package is created, click **Download** on each package to download the package to your local machine.

Move to production

On successful verification of content on the test server, start the process of moving content and configurations from the test server to the production server. First you should start with uploading the configuration from the test server onto the production server, followed by the content.

Move configurations

Perform the following steps to move the configurations on to the production server:

1) Disable the DAM Update Asset workflow.

2) Bring the package file containing configurations from your test server to the production server.

3) Open CRX Package Manager and click **Upload Package.**

4) Select the package that contains the configurations.

5) Click **OK**.

   All configurations from the test server are restored on the production server.

6) Verify the restored configurations. You can check your topic templates, DITA-OT profiles, and other configurations.

Move content

Content migration from one AEM instance to another could fall under the following two scenarios:

• When the test server and the production server are identical. Or, you are moving content from one fresh server to another fresh server.

• When you are moving content onto a server that already has some content and you want to move new content on that server.

Transferring files on identical or fresh servers

In case of the first scenario, you can simply copy-and-paste the `crx-quickstart` folder from the test server onto the production server. This is a quick and simple method of migrating not only your content, but the configurations as well.

Transferring files on different servers

In case the source and destination servers are not identical, then you have two ways of taking you content from one server to another:

Using the VLT Tool

   The VLT Tool is a flexible and powerful tool that allows you to transfer files from one system to another system in batches and with sufficient time gaps in-between two batched. For more information about the VLT Tool and how to use it, see the How to use the VLT Tool article.

If you are aware of scripting, then you can also create a script to move your files without any intervention.

Primarily, the following command is used to transfer files from one instance of AEM to another:

```
rcp -b <batch_size> -t <throttle_in_seconds> -r -u -n <source_path>
<destination_path>
```

For more options supported by the `rcp` command, see the **Rcp** command description in the How to use the VLT Tool article.

Using the content package

In case the source and destination servers are different and the content that you want to transfer is less than 500 topics, then you can use a content package. Perform the following steps to bring your content package from the one AEM server to another:

1) Bring the package file containing the content from the source to the destination server.
2) Open CRX Package Manager and click **Upload Package.**
3) Select the package that contains your content. If you are uploading multiple versions (or snapshots) of your content, then start with the first version and move to the latest version of the content.
4) Click **OK**.

   Your content from the source server is restored on the destination server.

5) Verify the restored content. You can check the DITA map or publish your content.

IMPORTANT: After importing the configurations and content, open the Configuration Manager (`http://<server-url>:<port>/system/console/configMgr`) and access the *com.adobe.fmdita.config.ConfigManager* bundle. Simply save the bundle without making any modifications. This refreshes your content and configuration changes