Configuring

Planning for retention of job and document data

Before you configure the Archive feature, plan the best way to set up repositories to hold the job and document data that you want to store.

You should identify:

- The files you need to store
- The properties you need to store
- The history information you need to store
- The points in the workflow where you need to store the job and document data
- How long you need to retain the job and document data
- Who should have access to the job and document data
- The system on which the job and document data should be stored
- What search properties you need to retrieve job and document data after they are stored in a repository
- What job property values you need to preserve for use when the job is retrieved from a repository

This information helps you decide how many repositories you need. You can then create the repositories and add **StoreInRepository** steps to your workflows.

Files

Each **StoreInRepository** step can store one file for each job in one repository. You can store any type of file in your repositories. For example, you can store PDF files, JDF job tickets, AFP files, PostScript files, or compressed files. You can store just property values or job history and no file.

If you have the PDF Document Support feature installed, you can retrieve data, including a PDF job, a document from inside a job, or information about a job, from a repository and view it. If you have the AFP Support feature installed, you can retrieve and view the same kinds of data for AFP jobs. You cannot view the contents of jobs or documents in other types of files, though you can view property values and job history if you have chosen to store that information.
You can create repositories based on who you receive files from or how you use the files in your company. For example:

- A print shop might store files for each customer account in its own repository.
- An in-house printing department might store files for each line of business in its own repository. One repository might store accounts receivable jobs, and another repository might store accounts payable jobs.
- A company might store files for each call center in its own repository. Call center employees have quick access to the files they need to support customers.

Properties

You can save the values of job and document properties with or without saving a job file or production history. You can also save the values of properties that are associated with a job, but are properties of other objects, and document data that is not defined as RICOH ProcessDirector document properties.

You can use job and document properties to retrieve a job, document, or information about the job from a repository and view the information. Information about the job can include properties associated with the job, such as the model of the printer used to print the job, and document data returned to RICOH ProcessDirector after an external program processes the documents in a job.

Storing properties takes much less space than storing jobs, so property repositories can be much smaller than repositories that store files.

You can specify properties to store based on how you use property information at your company. For example:
- A company might store the customer name, account number, and other information on each document printed as well as the number of pages in the document and the model of the printer used to print the job. Storing job files containing the actual documents might not be necessary.
- An in-house printing department might store job files, job and document properties needed to retrieve the documents in each job file, and postal processing data (such as change of address information) returned by third-party mail software. Although the postal processing data is not defined as RICOH ProcessDirector document properties, department personnel can view the data when they retrieve a document from the repository.

Production History

You can save the production history of a job, with or without saving a job file. You can set up history record notifications to capture varying levels of detail about job processing and then set the **StoreInRepository** step to store history records. For example, you can set up your notifications to record every state change for every step in the workflow that processes the job. Or you can limit what you record by choosing a particular state change, such as when the job state changes to **Complete** in the **InsertJobs** step or in the **PrintJobs** step.
Storing job history takes less space than storing jobs or documents, so job history repositories can be much smaller than repositories that store files or files and job history together.

You can create history record notifications based on how you use job history in your company. For example:

- A print shop might want to be able to tell customers precisely when their jobs are received, printed, and ready for pickup.
- A company might want to create an audit trail for all processing performed for specific customers whose jobs are time-sensitive.

Workflows

You can save files, job history, or both in a repository by adding a step based on the **StoreInRepository** step template at any point in the workflow. You can have multiple **StoreInRepository** steps in one workflow.

If you want to save jobs at different points in the same workflow, you can create a separate repository for each step. For example, you might store jobs in one repository before you process them so you have a record of the files that you received from customers. You might store jobs in another repository after you print them with added production information such as postal barcodes. The history records stored by each step contain only the job state changes recorded at the time each step ran. When you search for a job from the repository that was processed by multiple **StoreInRepository** steps, more than one result is returned. Each result contains only the history information known at the time the step ran.

> **⊕ Note:** Some steps that you can add to your workflows only work at the job level; they are not aware of the documents inside the jobs. Some of those steps, such as **ReversePDFPageOrder**, change the order of pages in the job, without consideration for the document properties that are associated with each page. When the pages are rearranged, the document properties remain associated with their original page locations, even though the content of that page is entirely different. Opening the job in the viewer and searching on document properties results in incorrect search results.
>
> For example, you receive a file with 50,000 pages from your customer. The workflow adds barcodes, sorts the job based on postal codes, then reverses the order of the pages so that the last page is printed first.
>
> If you store the job in a repository after the **IdentifyDocuments**, **IdentifyPDFDocuments**, **BuildAFPFromDocuments**, or **BuildPDFFromDocuments** steps, and before the **ReversePDFPageOrder** or **ReverseOutputOrder** steps, you can retrieve the correct documents in a job using document properties as search criteria. If you store the job in a repository after the **ReversePDFPageOrder** or **ReverseOutputOrder** steps, searching for a document in a job produces unreliable results.

After you add a **StoreInRepository** step to a workflow, you can adjust the retention period in the **RetainCompletedJobs** step. For example, you have set a long retention period in the **RetainCompletedJobs** step because you might need to reprint a job or document. You can reduce the retention period because now you can reprint from the repository.

Retention periods

If you need to retain different jobs or job history for different periods, create a separate repository for each retention period.

When you first set up a repository, you might want to set a short retention period so you can test the information specified in the **StoreInRepository** step. Set a one or two day retention period and send a sample job through the workflow. Then you can check the repository to be sure you stored the correct file, along with its properties and history information. Also verify that the history record notifications you set up are capturing the history information you expect.

You cannot delete jobs or documents from a repository or change the properties stored with them, so you must make sure that all of the information you expect to be retrieved is available. After you complete your testing, you can use the **Change Retention Period** action on the Repositories page to set the value you want to use for your production work.

Access

If you want to control access to a repository, you can set a value for the **Repository location** property. Only users who have access to the location set for the repository can find the documents and jobs stored in it. If different groups of users should have access to different stored jobs and documents, create a separate repository for each group.

System storage

If you want to store files in folders at different places on your computer system, create a separate repository for each folder. You can create a repository on the local system or on a mounted drive anywhere on your network.

Search properties

You can search for and retrieve jobs, documents, and their history information from a repository based on job or document properties stored in the archive.

For each **StoreInRepository** step in each workflow, identify the properties that you need to search for stored jobs, documents, or job history. For example, you might search by the **Job name** and **Customer name** job properties. You might search for a document by the **Member number** document property.

> **⊕ Note:** The Archive feature provides a large number of job and document properties that you can use to search for jobs and documents in a repository. Examine the values for the **Job properties to store** and **Document properties to store** properties on the **StoreInRepository** step template. If you need custom document properties, you can create them when you enhance PDF or AFP files for use with Archive.

In a step based on the **StoreInRepository** step template, you specify the repository to store the job in and the job and document properties used to search for jobs and documents in the repository. You can specify the same repository for multiple **StoreInRepository** steps. You can specify different search properties for each step.

If you use different search properties for different jobs, consider creating a separate repository for each set of jobs that use different search properties.

We recommend these best practices:

- Specify the same search properties for each step that stores jobs in the same repository.
- Store jobs with different search properties in different repositories.

- Choose properties carefully because after a job or document is stored in the repository, you cannot change its properties or remove the job or document from the repository.

When the search properties in a repository are consistent, RICOH ProcessDirector displays search results faster and data is simpler for users to find.

For example, repository A has 15 search properties. The properties are the same for each step that stores jobs in the repository. Repository B has 10 search properties. Some jobs stored in repository B use five of the ten properties. Other jobs stored in repository B use three of the properties. RICOH ProcessDirector displays search results faster for Repository A. Because the properties used to search for some jobs are different from the properties used to search for other jobs, users searching Repository B can become confused about which properties to specify. If a job or document does not have a value for the property that a user searches for, the job or document is not returned in the search results.

Search speed
How you set up repositories can affect the speed of repository searches. Keep these points in mind:
- Searches are faster if a repository has only jobs that were stored with the same search properties than if the repository has jobs that were stored with a mixture of different search properties.
- If a repository has only jobs that were stored with the same search properties, searches are faster if fewer search properties were stored.
- Smaller repositories are faster to search than large repositories.
- Search speed is not affected if you store data files along with properties and history information.

Preservation of job property values
You can store the values of job properties set when a job is processed to preserve them for use when the job (or a document in the job) is retrieved from a repository and submitted as a new job. For example, if a job has custom job properties whose values indicate special processing requirements, you can store those values in an override properties file and submit them with the job or document when you retrieve it from the repository.

Parent topic: [Setting up the Archive feature](#)