

Automation Engine

PDF Normalization

Contents

1. Introduction.....	3
2. What Makes a Normalized PDF Special?.....	4
3. Which Tasks Require PDF Normalization?.....	5
4. How do PDF Tasks Work with Normalized PDF Files?.....	6
5. How Do Normalized PDF Tasks Work With Regular PDF Files?.....	7
6. PDF Normalization and Step and Repeat.....	10

1. Introduction

This document explains the improvements in Automation Engine 12 with regards to the normalization of PDF documents.

Document Conventions

In the remainder of this document, PDF files that have not gone through normalization will be referred to as “PDF”. PDF files that have been normalized will be referred to as “Normalized PDF”.

Normalization in Automation Engine 10

In Automation Engine 10 and older versions, PDF files had to be normalized at the start of a workflow before any further processing could take place.

With the release of Automation Engine, 10.1 new PDF processing tasks were introduced that can handle incoming PDF files without the need for normalization. However, these tasks could not process normalized PDF files. On top of this the PDFs produced by these tasks needed to be normalized before they can be handled by other tasks. As a consequence, it was not straightforward to build workflows that used a combination of these tasks.

Improvements in Automation Engine 12

In Automation Engine 12, the collection of PDF processing tasks that do not require normalization was extended. To avoid that the user needs to carefully consider which tasks require normalized PDF as input when building workflows 2 features were introduced in Automation Engine 12:

- Tasks that do not require normalized PDF (the new ones added in Automation Engine 12 and the ones that were added in Automation Engine 10.1) can now also handle normalized PDF, and keep the files normalized.
- Tasks that require PDF files to be normalized will now do this automatically without the need to introduce explicit normalization tasks in a workflow.

2. What Makes a Normalized PDF Special?

Normalized PDFs are to a large extent regular PDF 1.6 files, except in a number of specific areas.

Overview

Normalized PDF differs from regular PDF 1.6 in the following areas:

- External references
- Color space
- Document level metadata
- Object level metadata

External references

A Normalized PDF is not always self-contained. It can refer to images or other PDFs that are not embedded in the PDF itself. Making use of references instead of embedding images and/or graphics can have several advantages in the workflow. The main advantage is that it allows for “late binding” e.g. an image can be color corrected. After the color correction all PDFs that refer to this image will automatically contain the corrected version of the image without the need to open and rewrite all those PDFs. Working with Normalized PDFs with external referenced can also lead to faster processing especially when the Normalized PDF refers to large amounts of external image data.

Although referring to external files is a standard PDF feature (e.g. used in the PDF/X-5), most PDF applications do not support this. To guarantee that a PDF can be processed correctly by PDF application (like Adobe Acrobat) it has to be made self-contained (embedding all the referred components like images). A self-contained Normalized PDF is 100% PDF compatible.

Color space

A Normalized PDF can only contain objects in the final printing color space. This final printing color space can have any number of separations but in most cases this will be CMYK and some spot colors. The PDF object model allows objects to be defined in RGB and CIE based color spaces. This is not possible in Normalized PDF. During normalization such objects are color converted to CMYK.

The fact that Normalized PDF can't contain objects in RGB or CIE based color spaces its biggest limitation. On the positive side it makes the workflow more predictable. It avoids late and often somewhat hidden color conversions e.g. in the RIP that can lead to problems in printing.

Document level metadata

A Normalized PDF contains XMP metadata. The schema of the XMP data is Esko specific, but the specification is public. The XMP metadata serves 2 purposes. It contains info about the file that can be extracted by 3th party applications and used e.g. for quality control or asset management. The XMP metadata are also used by Esko applications like Automation Engine to retrieve information from a PDF (e.g. the number of separations) in an efficient way.

Object level metadata

A Normalized PDF contains metadata for specific objects like barcodes. Object level metadata are added to allow modification of those objects using an Esko application. Object level metadata are stored using a proprietary mechanism and format.

3. Which Tasks Require PDF Normalization?

The following is an exhaustive list of the Automation Engine tasks that require normalization of PDF documents.

The table below lists the tasks that require normalization (Normalized PDF tasks) and the tasks that do not require normalization (PDF tasks). The exact list of tasks available on your Automation Engine server depends on the licenses you have, so some tasks in the list below may not be visible in your Automation Engine Pilot.

Normalized PDF Tasks	PDF Tasks
Add Ink Eaters	Adjust PDF Screening
Check Job Parameters	Compare PDF
Check Print Rules (PRC)	Create PDF File from RunList
Convert CMYK Colors	Extract Inks
Convert Colors	Optimize PDF Document
Create Design (DesignWizard)	Optimize PDF Separations
Create PAF/JPG/XML (LinkEdge)	Preflight with Pitstop
Create Report (ReportMaker)	Process PDF Spreads
Enrich Black	Resize PDF for Output
Export to 3D	Resolve OPI
Export to ArtPro File	Split Pages
Export to EPS/DCS File	Version PDF
Export to PDF File	
Export to PostScript File	
Export to VRML File	
FastVariants (all tasks)	
Optimize and Clean (PSFix)	
Outline Fonts	
PowerTrapper	
Prepare Station	
RIP (all tasks)	
Step & Repeat (all tasks)	
Trap (all tasks)	

4. How do PDF Tasks Work with Normalized PDF Files?

The example workflow described here takes a 1-up, removes the separation called “Cut” and traps the file. It can be used with both Normalized and regular PDF documents. In this topic, we take a look at how Automation Engine processes a Normalized PDF input file in this workflow containing PDF tasks.



Task	Task Type	Description
Optimize PDF Separations	PDF task	<p>The task checks whether the Normalized PDF input file contains any external references to images or other Normalized PDFs:</p> <ul style="list-style-type: none"> if there are external references, the Optimize PDF Separations task checks if the modifications it is instructed to make will affect the external references. If any of the external references are affected, the task will first make the Normalized PDF input file self-contained (by embedding all the external references). if there are no external references, or if the external references are not affected by the Optimize PDF Separations task will process the Normalized PDF directly. <p>The task also updates the XMP document metadata.</p>
Trap - Prepare Trap Layer	Normalized PDF task	The task traps the Normalized PDF file.

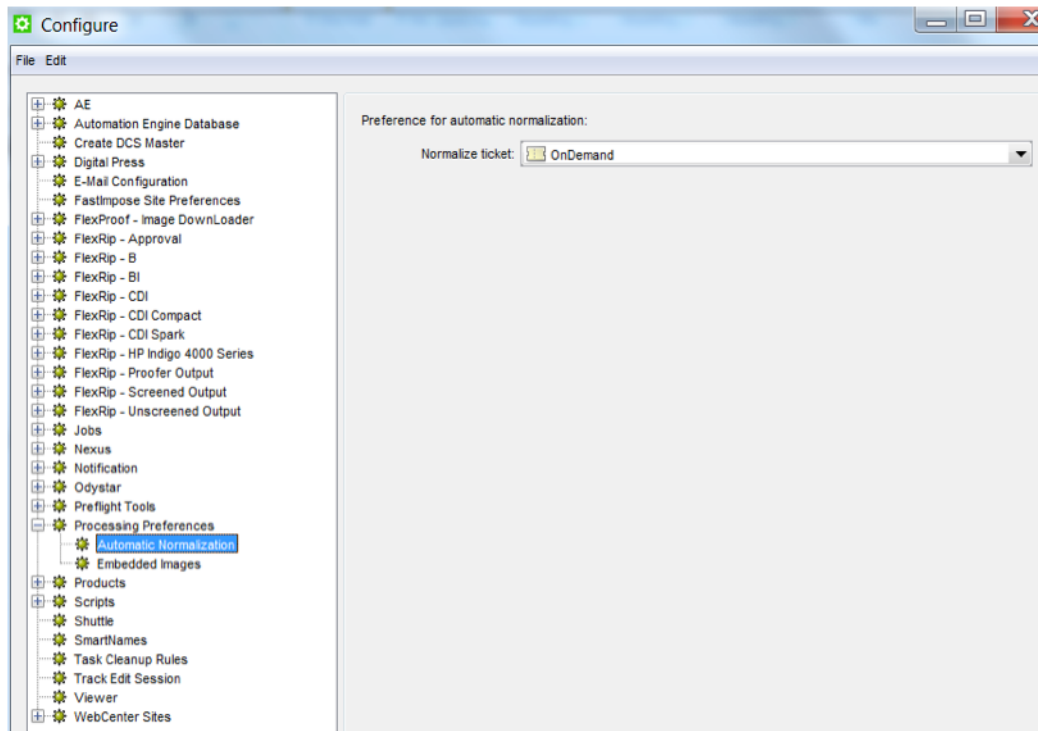
5. How Do Normalized PDF Tasks Work With Regular PDF Files?

In this topic, we take a look at how Automation Engine processes a regular PDF input file in our sample workflow containing a Normalized PDF task.

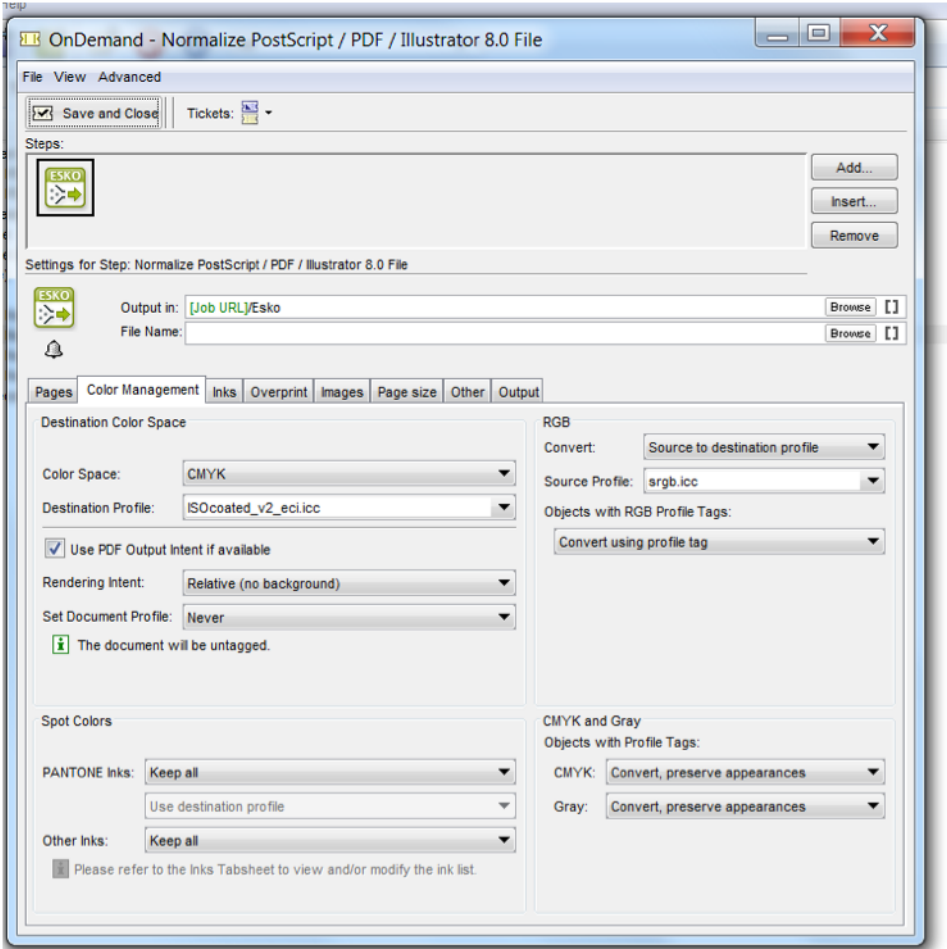


Task	Task Type	Description
Optimize PDF Separations	PDF task	The PDF is processed directly, because Optimize PDF Separations is a PDF task.
Trap - Prepare Trap Layer	Normalized PDF task	The task detects that the input file is a regular PDF document, and automatically normalizes it. After that, the file is trapped and a self-contained Normalized PDF is produced.

The ticket used for Automatic Normalization can be configured using the Configure window in the Automation Engine Pilot (choose **Tools > Configure** and navigate to **Processing Preferences > Automatic Normalization**).



Automation Engine 12 is pre-configured to use the “Normalize PostScript/PDF/Illustrator 8.0 File” ticket called “OnDemand”. From this ticket only the settings in the “Color Management” tab are taken into account. All other settings of the auto normalization are forced to fixed default settings. For example, automatic normalization will always embed all images.



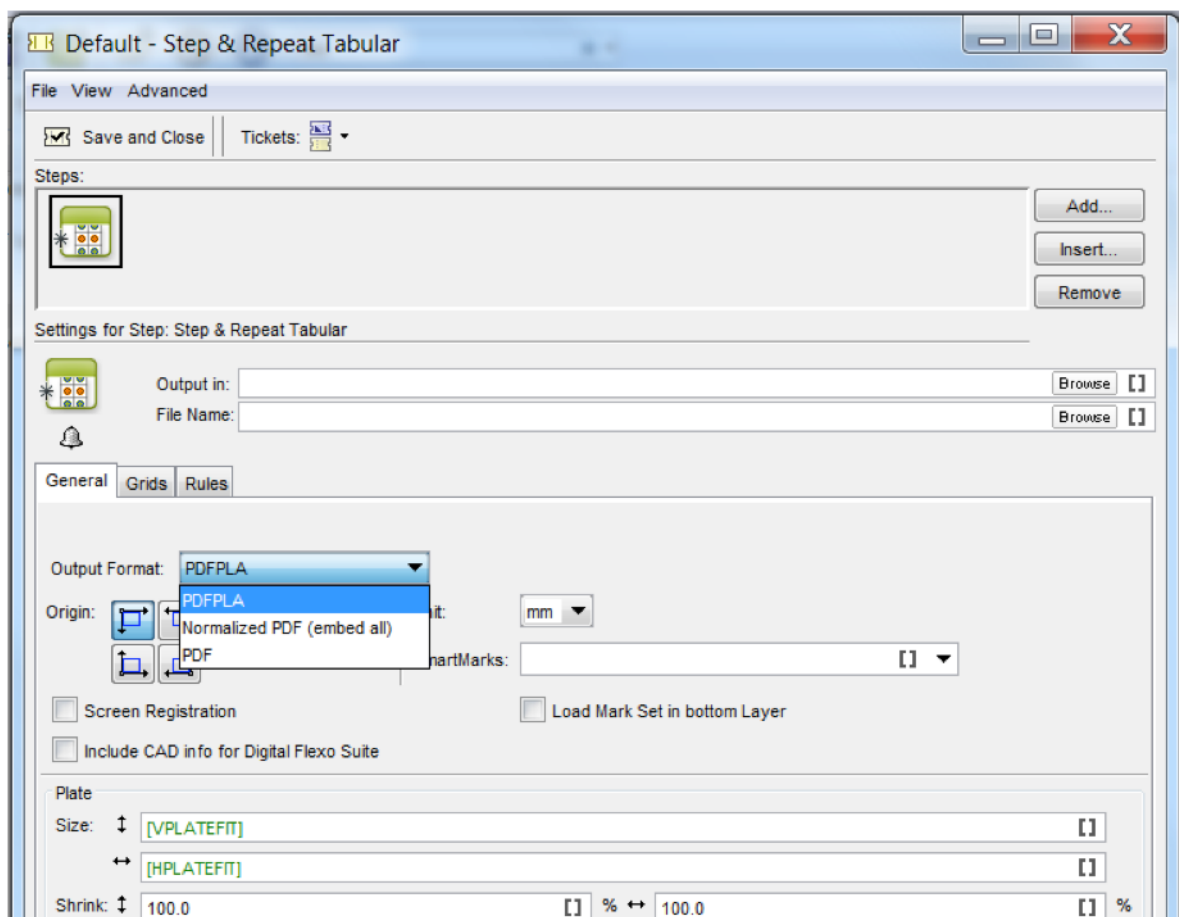
6. PDF Normalization and Step and Repeat

The step and repeat tasks in Automation Engine 12 can handle PDF, Normalized PDF or a mix of both at the input. The resulting output depends on a setting in the step and repeat ticket.

Output formats

Depending on your choice in the **Output Format** list of the **Step & Repeat** ticket, the format of the resulting document can be:

- PDFPLA
- a self-contained Normalized PDF
- a regular PDF



PDFPLA

When you choose this option, the output of the step and repeat task will be a PDFPLA file. A PDFPLA is a Normalized PDF with external references to the 1-ups. It can be loaded and modified in Plato. It can be RIPped directly by FlexRip. To send it to a non-Esko system, the PDFPLA needs to be exported to a PDF using the **Export PDF** task.

Note:

PDFPLA output can only be generated if all the inputs are Normalized PDFs. The task will fail if a 1-up at the input is not normalized on beforehand.

This format is the recommended output format when the primary purpose of the step and repeat file is to send it to FlexRip.

Normalized PDF (embed all)

When you choose this option, the output of the step and repeat task is a self-contained Normalized PDF. The 1-ups are embedded in the PDF using PDF forms so that the file size remains compact. The file can not be modified in Plato. It can be RIPPed directly by the FlexRip. It can also be sent to a non-Esko system.

Normalized PDF inputs will be embedded in the output. PDF inputs will be auto normalized and then embedded in the output.

This format is recommended when further processing is needed on the output, for example by the Optimize PDF Separations task before sending it to FlexRip.

PDF

When you choose this option, the output of the step and repeat task is a PDF. The 1-ups are embedded in the PDF using PDF forms so that the file size remains compact. The file can not be modified in Plato. FlexRip can not RIP this file directly: it will first automatically normalize the step and repeat file. The file can be sent to a non Esko system.

Normalized PDF inputs will be embedded in the output. PDF inputs will be embedded in the PDF "as is".

This format is recommended when the primary purpose of the step and repeat file is to send it to a third-party, non-Esko RIP.