

## IBM® Rational® Rhapsody® Gateway Add On



## PDF Coupling Notes



***Rhapsody<sup>®</sup>***

**IBM<sup>®</sup> Rational<sup>®</sup> Rhapsody<sup>®</sup>  
Gateway Add On**

**PDF Coupling Notes**



## **License Agreement**

No part of this publication may be reproduced, transmitted, stored in a retrieval system, nor translated into any human or computer language, in any form or by any means, electronic, mechanical, magnetic, optical, chemical, manual or otherwise, without the prior written permission of the copyright owner.

The information in this publication is subject to change without notice, and Dassault Systèmes and its affiliates assume no responsibility for any errors which may appear herein. No warranties, either expressed or implied, are made regarding Rhapsody software and its fitness for any particular purpose.

## **Trademarks**

Reqtify is a registered trademark of Dassault Systèmes or its affiliates in the US and/or other countries.

Rhapsody Gateway, IBM, the IBM logo, DOORS and Rhapsody are trademarks or registered trademarks of IBM Corporation.

All other product or company names mentioned herein may be trademarks or registered trademarks of their respective owners.

© Copyright 2001-2011 Dassault Systèmes. All rights reserved.

# Contents

---

<b>Contents .....</b>	<b>5</b>
<b>Introduction .....</b>	<b>7</b>
<b>PDF Analysis.....</b>	<b>9</b>
Selecting the PDF File .....	9
Specific Parameters of a PDF based Type .....	10
Adding a Cover.....	12
<b>Requirements Traceability .....</b>	<b>13</b>
Traceability Information in PDF.....	13
Intermediate File .....	13
Basics .....	14
Tables in PDF Documents .....	15
Capturing PDF Annotations .....	15



# Introduction

---

This technical note describes how to capture the traceability information of requirements from a PDF document using Rhapsody Gateway.

Some standard operations, such as navigating in an Acrobat document from the Rhapsody Gateway tool, require the Acrobat PDF tool to be installed.

## Prerequisite

Java 1.5 or higher version shall be installed in the computer. The use of Java 1.6 is recommended.

Refer to the *Customization Guide* and run the Coupling example for more information about what is discussed in this technical note.



# PDF Analysis

---

This chapter describes how Rhapsody Gateway captures information items from PDF and how these information items are analyzed to be used as requirements, or references to higher level requirements.

This chapter covers the following topics:

- ◆ Getting the information items from PDF.
- ◆ Configuring Rhapsody Gateway according to the traceability information in PDF.



## Note



---

The PDF format analysis is very limited, because sometimes the PDF format is encrypted and/or protected. Rhapsody Gateway can only work within these limitations and protections. For example, it will not be able to analyze an encrypted PDF.

## Selecting the PDF File

To insert a PDF file, follow these steps:


1. In the Project editor, click on the **Add a document**  button to insert a new document into the Traceability Description Area.
2. Select a **Type of Analysis** based on PDF.
3. Click on the **File or Directory** field then on the **Browse**  button to select a PDF file (with the pdf extension) from the Rhapsody Gateway selection window.

Details   Modification Files   Covers			
	Name	Type of Analysis	File or Directory   lg
	Requirements	 PDF	PDF Demo.pdf

## Specific Parameters of a PDF based Type

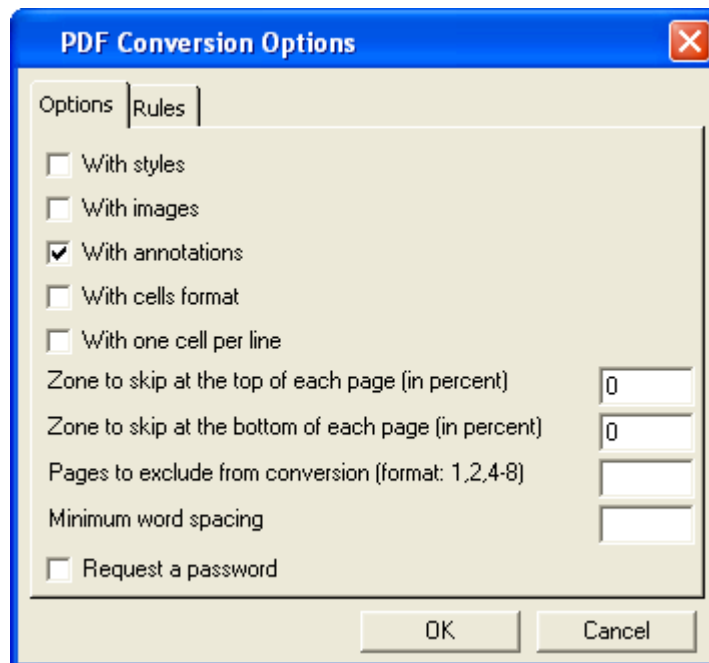
When you select a type of analysis defined to capture information from a PDF (either a default type or a customized one), you have the opportunity to define additional parameters:

- ◆ **Options**—allows accessing to a window presenting the PDF file options when clicking in the **Value** field.

Details   Modification Files   Covers								
Name	Type of Analysis	File or Directory	lg...	In...	Bl...	Variable	Value	
Requirements	PDF	PDF Demo.pdf	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Options		

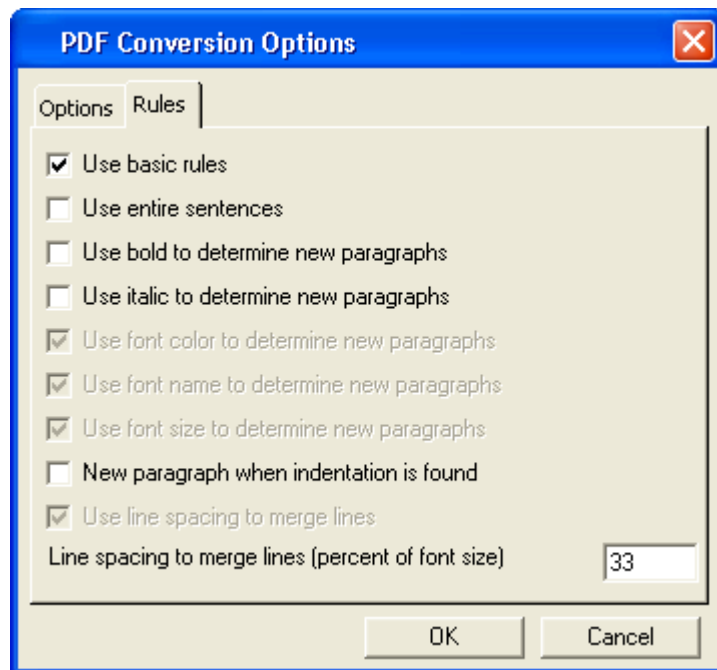
The setting window contents two panes:

- The **Options** pane contains the following parameters:



- **With styles**—allows extracting styles information from the PDF document.
- **With images**—allows extracting images from the PDF document.
- **With annotations**—allows extracting the PDF annotations. If this checkmark is not checked, PDF notes will not be taken into account. See the *Capturing PDF Annotations* section.
- **With cells format**—allows extracting the cells. See the *Tables in PDF Documents* section.
- **With one cell per line**—for tables, prints each cell in a separate line.

- **Zone to skip at the top of each page**—defines the height (in percent of page height) to skip from analysis at the beginning of each page.
  - **Zone to skip at the bottom of each page**—defines the height (in percent of page height) to skip from analysis at the end of each page.
  - **Pages to exclude from conversion**—Defines the number of the pages to exclude from the analysis (ex: 1;2;4-8).
  - **Minimum word spacing**—Defines the minimum space which should exist between two letters to identify a partition between two words.
  - **Request a password**—When this option is checked, the user will be prompted to enter a password each time he will convert the document when the PDF document is protected from read.
- The **Rules** pane contains a list of rules to choose:



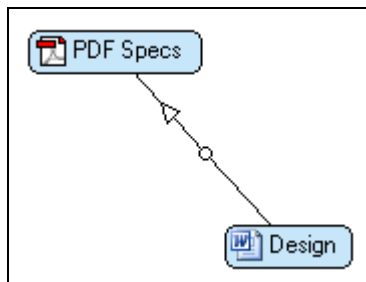
- **Use basic rules**—paragraphs are generated with font name, font color, font size and leading Rules. If this rule is checked, some other rules are checked by default.
- **Use entire sentences**—when a line is significantly shorter than the next line, this means it ends a paragraph, so a new paragraph is created for the next line.
- **Use bold to determine new paragraphs**—if bold status differs from a line to another, this begins a new paragraph.
- **Use italic to determine new paragraphs**—if italic font status differs from a line to another, this begins a new paragraph.
- **Use font color to determine new paragraphs**—if font color status is different from a line to another, this begins a new paragraph.

- **Use font name to determine new paragraphs**—if font name status is different from a line to another, this begins a new paragraph.
- **Use font size to determine new paragraphs**—if font size status is different from a line to another, this begins a new paragraph.
- **New paragraph when indentation is found**—if an indentation is found, this begins a new paragraph.
- **Use line spacing to merge lines**—lines spacing between two lines are used to merge them.
- **Line spacing to merge lines**—defines the height in percent of font height between two lines to merge them.

## Adding a Cover

To define traceability, add a cover between the covering document and the covered document.

The Rhapsody Gateway project looks like the following:



Now the traceability analysis of the project can be performed.

# Requirements Traceability

---

There are many ways to enter traceability information in a document before converting it to PDF.

## Traceability Information in PDF

Default configuration of the PDF type of analysis considers:

- ◆ Requirements:
  - are captured when they use the syntax `<ID> : <requirement label>`. `<ID>` is the requirement ID, and must be written REQnn (REQ followed by digits). `<requirement label>` is the “title” of your requirement, usually written using natural language. Use a colon to separate the ID from the Label. If you do not use this colon, all information will be captured as an ID. The label, however, is optional.
- ◆ References:
  - Syntax is `[Covers: <requirements>]`. Use the string `[Covers:...]`, `<requirements>` are your requirement ID(s). If you list several requirement IDs, insert , or ; between them. For example, use `[Covers: REQ1]` or `[Covers: REQ1,REQ2,REQ3]`
  - You can create references that will cover your higher level document requirements in a PDF document already managed by Rhapsody Gateway. In order to do so, select a requirement in this document, right-click the requirement and select Copy For > Covers Reference in the contextual menu.

## Intermediate File

Rhapsody Gateway creates an intermediate file called `<document_name>.txt`, containing the information extracted from the PDF document.

You can refer to the PDF file *PDF demo.pdf* used in the Coupling PDF demo example to analyze how the converter works.

This intermediate file is usually deleted at the end of the conversion process. However, you can keep this helpful intermediate file available, for training or customization purpose, by activating the option in the Project Editor.

Details   Modification Files   Covers					
	Name	Type of Analysis	File or Directory	Ignor...	Intermediate file
	Requirements	PDF	PDF Demo.pdf	<input type="checkbox"/>	<input checked="" type="checkbox"/>

## Basics

Consider for instance the following PDF document extract:

<p><b>2 DEMO REQUIREMENTS</b></p> <p><b>REQ1 : Capture</b></p> <p>&lt;&lt;The tool shall be able to capture semi-automatic document and/or in a model. "Semi-automatic beforehand by the user or another dedicated tool.&gt;</p> <p><b>Allocated to : John Doe</b> <b>Priority : High</b></p> <p><b>REQ2 : Update information when source changes</b></p> <p>&lt;&lt;The tool shall take into account the successive update automatically the traceability information p</p> <p><b>Allocated to: Tim Jones</b> <b>Priority : High</b></p> <p><b>REQ3 : Multiple requirements definition</b></p> <p>&lt;&lt;The tool shall be able to consider several requi example, if Safety requirements, Functional requi</p>
---

The intermediate file is then created in the **intermediate** sub-directory of the project directory.

```

2 DEMO REQUIREMENTS
REQ1 : Capture
<<The tool shall be able to capture semi-automatically the requiremen
Allocated to : John Doe  Priority : High
REQ2 : Update information when source changes
<<The tool shall take into account the successive versions of the do
Allocated to: Tim Jones  Priority : High
REQ3 : Multiple requirements definition
<<The tool shall be able to consider several requirements definition.

```

## Tables in PDF Documents

For tables the intermediate file contains:

```
|<column number> Cell text
```

For several paragraphs in the same column, all paragraphs are presented on the same line before switching to the next column. For example, if you have 2 paragraphs in column 1:

```
|1 1st paragraph|1 2nd paragraph|2 Second column...
```

## Capturing PDF Annotations

Traceability information can be added in annotations. To capture annotations, the **With annotations** option needs to be checked and a customized definition of requirement needs to be created.

In the customized type, a regular expression corresponding to this capture can be:

```
^Annot [^/]+/\s+([^\s/]+) /\s+([^\s/]+)
```

- ◆ The first parentheses group captures the requirement ID.
- ◆ The second parentheses group captures the requirement text, if **Text** is selected for **Field2** in the contextual menu of the requirement field. For more details, refer to the *Defining Requirement* section in the *Customization Guide*.

The corresponding intermediate file contains the following lines:

```
Annot_<username>_<noteGUID>/<space><contents of  
annotation>/<space><highlighted PDF text>/End_annotation
```