



# Report Mining

Brian Leach Consulting Limited

<http://www.brianleach.co.uk>

Brian Leach Consulting Limited makes no warranty of any kind with regard to the material contained in this manual, including but not limited to the implied warranties or merchantability and fitness for a particular purpose.

The information contained in this document is subject to change without notice.

This document contains proprietary information that is protected by copyright. All rights are reserved. It may not be photocopied, reproduced, or translated, in whole or in part, without the prior express written consent of Brian Leach Consulting Limited.

Copyright 2009 Brian Leach Consulting Limited. All rights reserved.

# Contents

Acknowledgments .....	4
Introduction : mvPDF Report Mining.....	5
Reporting Mining in mvPDF .....	6
Using the Report Mining Feature.....	7
Defining the Report Map .....	9
Starting the Report Zone Editor .....	10
Loading a Report .....	11
Defining the Report Zones .....	12
Adding a Zone .....	12
Changing an Anchor .....	14
Adding a Simple Report Element.....	14
Adding Repeating Elements .....	20
Repeats and Text Wrapping .....	22
Adding Dead Zones .....	23
Saving a Report Map .....	24
Loading a Report Map .....	24
Running the Report Mining.....	25
Report Mining to a Work File .....	25
Report Mining to a Merge Form .....	25

## Acknowledgments

---

UniVerse, UniData and UniObjects are registered trademarks of IBM U2.

PICK and Pick Systems are registered trademarks of Raining Data Corporation.

Access Query Language, AQL, D3 and FlashBASIC are trademarks of Raining Data Corporation.

UNIX is a registered trademark in the United States of America and other countries, licensed exclusively through X/Open Company Ltd.

Microsoft, Windows, Windows NT, Windows 2000 and MS-DOS are trademarks of Microsoft Corporation. ActiveX, JScript, Visual Studio, and Visual Basic are either registered trademarks or trademarks of Microsoft Corporation.

Adobe and Acrobat are trademarks of Adobe Systems Incorporated.

All other company or product names mentioned are trademarks or registered trademarks of their respective trademark holders.

The example companies, organizations, products and domain names depicted herein are fictitious. No association with any real company, organization, and product or domain name is intended or should be inferred.

## Introduction : mvPDF Report Mining

---

Any MultiValue application of any standing is likely to include a wide variety of reports. These reports may be in different formats and may be created by different processes: Basic programs, enquiry listings or fourth generation languages like SB+.

However these reports have been created, they represent an important investment in time and business knowledge. In many ways, for the majority of users empowered with critical decision making, the reports **are** the application.

The problems with reporting are often perceptual. Reports built from MultiValue databases are often visually unappealing. If they are customer facing – invoices, statements, despatch notes – they may present an image of the company that is out of step with the times.

At the same time, clients demand new forms of output: emailed documents to replace hard copy printouts. The business may demand new access to information locked in the system, to feed into spreadsheets and analytical tools.

Often this may be a spur to purchase new reporting tools. But however glossy and functional these tools may seem, the technology is only part of the answer. Your chosen reporting tool may offer a wide range of exports and reporting features, but it is only as good as the investments you can make in helping it to uncover an understanding of your application. In these credit-crunch times, organizations are looking to 'up their game' in presentation and competitive advantage, but few have the resources to spend on deploying new reporting suites.

### **Report Mining offers a different perspective.**

Report Mining looks on your existing reports as a vital resource. True, they may look old and may not offer the dynamic operations offered by a spreadsheet: but they represent vital information drawn from your system, often refined over many years and through deep business knowledge.

Report Mining aims to reuse these reports, but in a new and exciting way. Why start again from scratch when you have the information you need at your fingertips?

Report Mining captures your reports, scans them and turns them into an information resource. The results may be re-presented or exported: the choice is yours. Report Mining lets you preserve your reporting investment whilst bringing it up to date with new business practices and presentation, at a fraction of the cost and risk of developing new reporting structures.

## Reporting Mining in mvPDF

mvPDF is a suite of document tools for MultiValue databases. mvPDF offers a range of features for assembling documents, ranging from a rich programmers API through data merging using templates and watermarks, to capture and conversion of existing reports.

Report Mining is a new and exciting extension to the capture and conversion features already found in mvPDF. With Report Mining, mvPDF is not restricted to converting the substance of your reports: now it can convert the information that is held within them.

mvPDF lets you map out the structure of your reports, using a visual report editor that identifies and categorizes the information present. Where reports contain many different sections and sets of information, they can be divided into Zones, based on placement or by locating elements such as headers and labels, allowing mvPDF to draw information from each section in a single operation. You can specify repeat patterns, text wrapping and dead zones to rip columns of related information from documents as wide ranging as enquiry listings, purchase orders or picking lists.

Name	Column Type	Column Value	Zone	Row Value	Extract Type	Extract Value	Trim Style	OConv	IConv
CLINAME	Column Position	4	A1	-4	Base on length	40	Both ends		
CLIADDR1	Column Position	4	A1	-3	Base on length	20	Both ends		
ACCOUNT	Column Position	104	A1	+0	Base on length	10	Both ends		
REFNO	Column Position	104	A1	+1	Base on length	10	Both ends		
PREPARED	Column Position	104	A1	+2	Base on length	12	Both ends		

Once mapped, mvPDF lets you run, prepare and rip your existing reports. The report content can be directed to a work file using a record layout of your choosing, ideal for exporting using tools like the free EXPORT utility and Conduit SQL database pipeline, both available from our website. Or you may augment the report data with supplementary information drawn through dictionary items, and combine these in an mvPDF document merge. Or finally, you may send the information directly through mvPDF without intervention to create new documents and printouts, choosing new layouts to present the information in a more vibrant and appealing style.

## Using the Report Mining Feature

The Report Mining feature is an extension of the way in which mvPDF captures and processes spooler jobs. Without report mining, you can run any routine that creates a regular database print following a **SET.PDF** command. This command switches the database spooler to capture mode and sets up the control information for processing the print job through mvPDF. This control information typically specifies the name of the PDF document to create, or the printer to which the reformatted job should be sent; the page size, font, authorship and security details and the name of a watermark page to apply.

Using watermarks can improve the visual appearance of your captured reports, in the same way as a PCL macro or the use of pre-printed stationery: but in regular capture mode mvPDF cannot structurally alter the presentation or layout of the report:

PETER MOFFET  
31 Coppice Lane  
HALESOWEN  
WEST MIDLANDS  
B 3YB


  

201	The Decline and Fall of the Roman Empire (Part	1	24.99	24.99
110	A Tale of Two Cities	1	64.57	64.57
45	The Hostile Hospital: Complete & Unabridged (Se	1	13.00	13.00
92	Making Divorce Work: In 9 Easy Steps	1	14.99	14.99
240	Voices of History	1	15.99	15.99
147	The Time Traveler's Wife	1	17.99	17.99
96	Cirque Du Freak: Complete & Unabridged (Saga of	1	12.99	12.99
106	Making Divorce Work: In 9 Easy Steps	1	13.99	13.99
194	The Twits	1	5.99	5.99

Invoice To:

PETER MOFFET  
31 Coppice Lane  
HALESOWEN  
WEST MIDLANDS  
B 3YB




# Invoice

Goods or Services:				
Part Number	Description	Qty	Price	Total
201	The Decline and Fall of the Roman Empire (Part	1	24.99	24.99
110	A Tale of Two Cities	1	64.57	64.57
45	The Hostile Hospital: Complete & Unabridged (Se	1	13.00	13.00
92	Making Divorce Work: In 9 Easy Steps	1	14.99	14.99
240	Voices of History	1	15.99	15.99
147	The Time Traveler's Wife	1	17.99	17.99
96	Cirque Du Freak: Complete & Unabridged (Saga of	1	12.99	12.99
106	Making Divorce Work: In 9 Easy Steps	1	13.99	13.99
194	The Twits	1	5.99	5.99
<b>ORDER TOTAL:</b>				<b>184.46</b>

With Report Mining, you can use the information in your reports in new ways. Instead of sending the report text directly to the mvPDF Server, the spooler capture post-processes the report to dig out and store the report content. You can then apply selected parts of the report directly to a regular mvPDF Merge Form:

Invoice To:  
PETER MOFFET  
31 Coppice Lane  
HALESOWEN  
WEST MIDLANDS  
B 3YB



# Invoice

Invoice Date : 25 MAY 2005

Goods or Services:		
Part Number	Description	Price
201	The Decline and Fall of the Roman Empire	24.99
110	A Tale of Two Cities	64.57
45	The Hostile Hospital: Complete & Unabridged	13.00
92	Making Divorce Work: In 9 Easy Steps	14.99
240	Voices of History	15.95
147	The Time Traveler's Wife	17.99
98	Cirque Du Freak: Complete & Unabridged	12.99
106	Making Divorce Work: In 9 Easy Steps	13.99
194	The Twits	5.99

ORDER TOTAL: 184.46

Or you can choose to use the extracted data in any way you wish:

```

ED PDFWORK 179.1
7 lines long.

----: p
0001: PETER MOFFET
0002: 31 Coppice LaneýHALESOWENýWEST MIDLANDSýB 3YB
0003: 25 MAY 2005
0004: 201ý110ý45ý92ý240ý147ý96ý106ý194
0005: The Decline and Fall of the Roman EmpireýA Tale of Two CitiesýThe Hostile
Hospital: Complete & UnabridgedýMaking Divorce Work: In 9 Easy StepsýVoi
ces of HistoryýThe Time Traveler's WifeýCirque Du Freak: Complete & Unabr
idgedýMaking Divorce Work: In 9 Easy StepsýThe Twits
0006: 24.99ý64.57ý13.00ý14.99ý15.95ý17.99ý12.99ý13.99ý5.99
0007: 184.46
Bottom at line 7.

```

## Defining the Report Map

---

In order to successfully perform mining operations on a report, you must tell mvPDF how to find and to organize the information that makes up the meaningful content of that report.

For example, consider an invoice print like the one below:

SALES INVOICE =====		
Invoice To:		
PETER MOFFET 31 Coppice Lane HALESOWEN WEST MIDLANDS B 3YB		
Invoice Date : 25 MAY 2005		
Item Title		Price
201	The Decline and Fall of the Roman Empire	24.99
110	A Tale of Two Cities	64.57
45	The Hostile Hospital: Complete & Unabridged	13.00
92	Making Divorce Work: In 9 Easy Steps	14.99
240	Voices of History	15.95
147	The Time Traveler's Wife	17.99
96	Cirque Du Freak: Complete & Unabridged	12.99
106	Making Divorce Work: In 9 Easy Steps	13.99
194	The Twits	5.99
		=====
		184.46

This contains some information that can be discarded, such as the SALES INVOICE header and the labels for Invoice To and Invoice Date. In amongst these you can see the important information: the name and address, invoice date, invoice total and the repeating lines defining the order details (item, title and price).

Each report that you wish to mine is accompanied by a Report Map. The Report Map tells mvPDF where to locate each of these pieces of information, how these should be named, and how the data will repeat. If the text in the details line wrap, you can also control the wrapping styles so that mvPDF can distinguish between long text and new value lines.

To assist you in building the map, mvPDF is supplied with a graphical Report Zone Editor.

### NOTE

The map itself is held in the PDFMAPS file. If you are using automated tools to create your report, such as a fourth generation language, you may also be able to populate the maps automatically from your report definitions. Please contact the author for details.

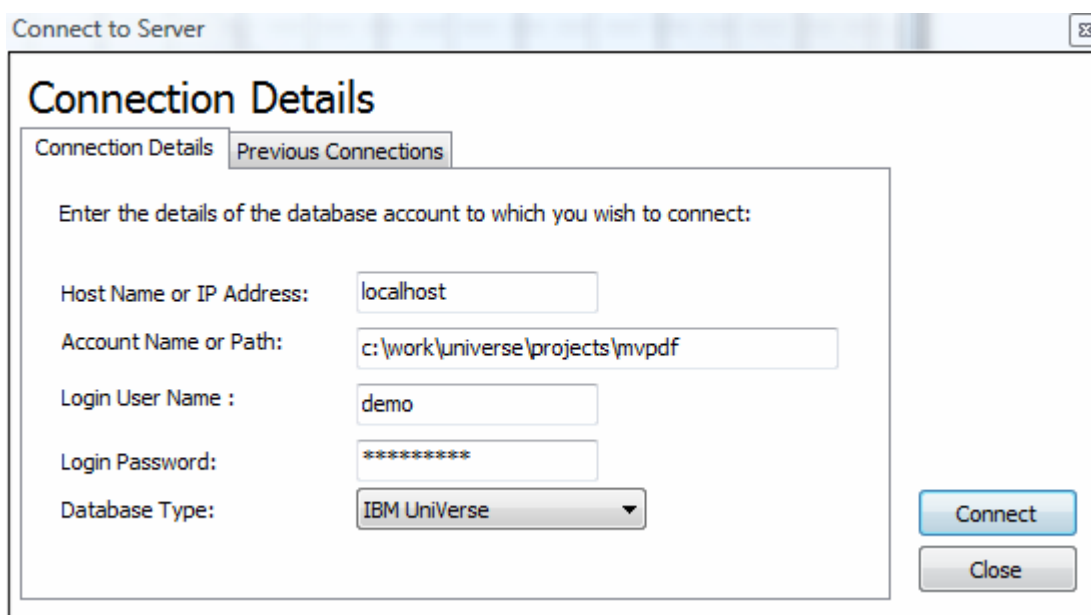
## Starting the Report Zone Editor

---

The Report Zone Editor is provided as part of the mvPDF setup. You can load the editor onto as many workstations as you wish.

To begin the editor, select Program Files ->BLC -> mvPDF ->mvPDF Report Zone Editor from your Start button.

Before you can edit a report map, you must connect the Report Zone Editor to the database account in which the report will be run. For this, select the **Connect** option from the **File** menu:



The screenshot shows a window titled "Connect to Server" with a close button in the top right corner. Inside the window, there is a section titled "Connection Details" with two tabs: "Connection Details" (selected) and "Previous Connections". Below the tabs, there is a text prompt: "Enter the details of the database account to which you wish to connect:". The form contains the following fields:

- Host Name or IP Address: localhost
- Account Name or Path: c:\work\universe\projects\mvpdf
- Login User Name : demo
- Login Password: \*\*\*\*\*
- Database Type: IBM UniVerse (dropdown menu)

At the bottom right of the form area, there are two buttons: "Connect" (highlighted in blue) and "Close".

Enter the connection details and click the Connect button to connect to the database. If the connection succeeds, the Report Zone Editor will store the connection details, with the exception of the password, for future use. You can select a previous connection from the second tab.

### NOTE

The Report Zone Editor uses a standard middleware to connect to the database, which must be loaded onto your PC. You must also install the middleware required for the editor to connect to your database:

For UniVerse and UniData, this is the UniObjects middleware located on the UniVerse/UniData client CD under the name UniDk or UniObjects Developers Kit.

For QM this is the QMClient library.

## Loading a Report

---

The Report Zone Editor consists of two panels. The upper panel displays the text of the report that you wish to mine, and the lower displays details of the report map.

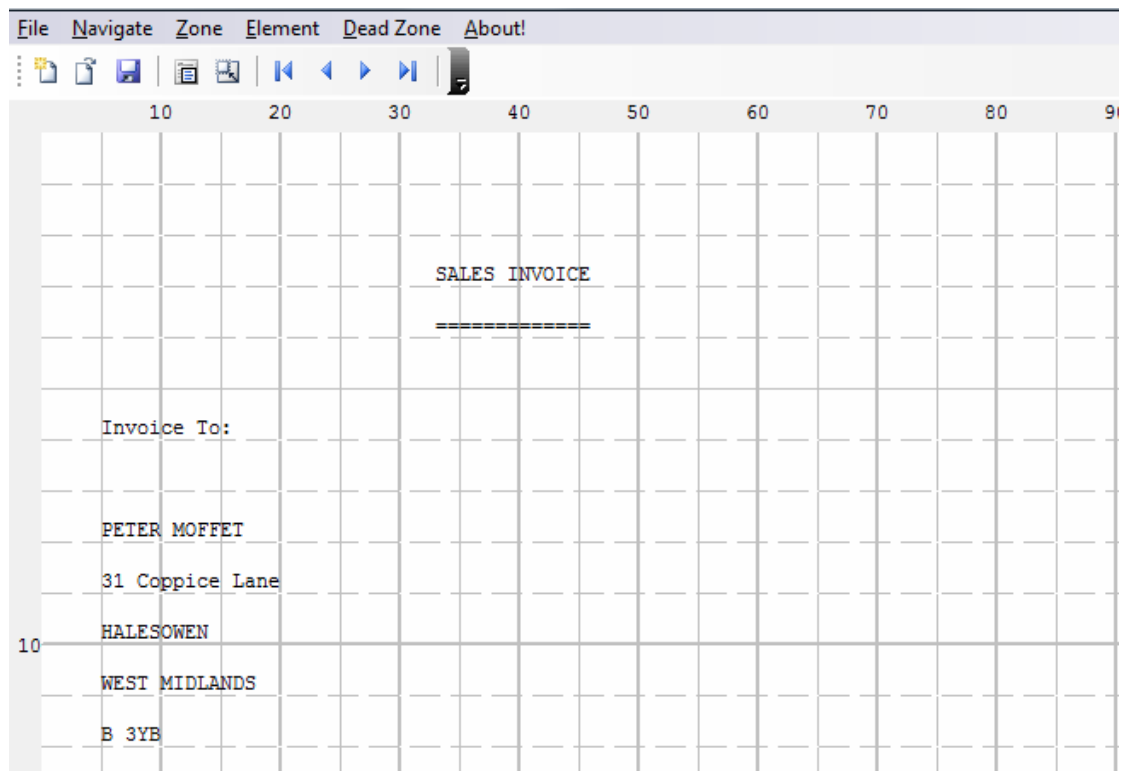
Once you have connected to your database, you will need to load your report. The Report Zone Editor requires that you have first run and captured your report, since it cannot predict the route that your application will take to generate the report, and does not know whether it is possible to create that report stand-alone.

If you do not know how to capture the report, you can use two utility commands in mvPDF:

**PDF.CAPTURE.START**  
**PDF.CAPTURE.STOP**

The **PDF.CAPTURE.START** should be issued before you run your report. This will capture any prints produced by your application. The **PDF.CAPTURE.STOP** should be run after your report to turn off capturing and to provide you the option to say where to save the captured report.

Once you have captured your report, you can load it from the File menu. The Report Zone Editor displays the report text:



The text is displayed with double line spacing to make the report elements that you will add more legible. You can use the navigation button on the tool bar to move up and down the report a screen at a time. If you have a wheel mouse, you can use this to move up and down a line at a time.

## Defining the Report Zones

---

Each Report Map is divided into a number of Zones. These zones then contain the various elements that you wish to extract.

Zones make it easy to identify the various parts of a report. Zones signal the row at which a set of information begins: this might be the top of a page, a particular label or the starting row of a set of repeating values. A simple report such as an enquiry listing may only consist of a single zone. A more complex report holding many different sets of information may consist of many zones.

There are two ways in which mvPDF can identify a Zone:

- A Zone may be fixed to the start of a page or the start of the report.
- A Zone may be variable and based on some specific text.

For example, if all of your pages have exactly the same layout, as for an enquiry listing, you can set a zone that will be anchored to the top of each page. Each element can then be identified by its absolute position on the page. This is the easiest option, but the more limited.

Often it is better to identify zones based on text. This allows the zone to begin anywhere in the report: for example, where you may have a variable number of detail lines appearing above it.

In the example Invoice report above, there are two Zones required to extract all the information you need. The first can be used to extract all of the information that begins at a specific location in the report: the NAME, ADDRESS, DATE, ITEM, TITLE and PRICE. It does not matter that the last three fields continue over multiple lines: the Zone only needs to identify where the first line of an element begins.

However, the final piece of information - the TOTAL – does not always start at the same point since it follows on from the variable detail lines above. Instead, you can set a Zone that will be identified by the bars running across between the last PRICE and the TOTAL (=====) and appearing at that specific column.

Learning to identify the report zones can be tricky at first, but can also make it very much simpler to lay out a report map.

## Adding a Zone

---

To add a Zone to your Report Map, you can select the Add option from the Zone menu or click the Add Zone button on the toolbar.

TIP:

If you are adding a zone based on some text, you should highlight that text in the report window before you add the zone. The Editor will feed your selection into the Zone Wizard, saving you the job of entering the details.

This displays the Zone Wizard:

The image shows a 'Report Zone' dialog box with the following fields and options:

- Zone Name:** A1
- Zone Location:**
  - Anchor Type:** Anchor to Text
  - Look for text:** Invoice To:
  - In this column: (leave blank for any):** 6
- Options:**
  - Start new Record
  - Clear all Anchors

Buttons: Cancel, Ok

**Zone Name:**

Each zone is identified by a unique name. The Report Editor will assign a default name of *A<sub>n</sub>* where *n* is the next zone number. If you are mapping a complex report you can change this to something more meaningful.

**Anchor Type:**

Each Zone begins at a given Anchor point. There are three types of anchors available:

- Text
- Top of Form
- Top of Report

All the elements in the zone are identified by row coordinates offset from the anchor point.

Text anchors are based on specific text found in the report: typically elements such as column headers, labels or sub total lines. The text to find must be specified exactly as it appears in the report, in the **Look for text** box below. If you want to refine the search, you can specify a starting column (from 1) at which mvPDF should look for the text.

If the Zone is Top of Form, each element will be identified by its absolute position on each page. These are reset for each page in turn.

If the Zone is Top of Report, each element is identified by absolute position in the report, irrespective of any page breaks in-between. This is only useful if each report deals with a single topic: e.g. where you have a single invoice in the report, and not multiple invoices each beginning on a separate page.

In the example above, the Anchors can both be set to Text. The first anchor could be identified by the SALES INVOICE header or by the Invoice To: text. The second will be identified by the subtotal line divider (=====) under the PRICE field.

### Start New Record:

As mvPDF mines the report, it places the text that it extracts into a work file. As a default it will create a single record holding all of the data, but for a large report in particular this is not desirable. You can create a new record each time a Zone has begun, or for each repeat in a set of detail lines.

If you are using a merge form, you should set the New Record to the zone that identifies where the merge form data will begin.

### Clear Anchors:

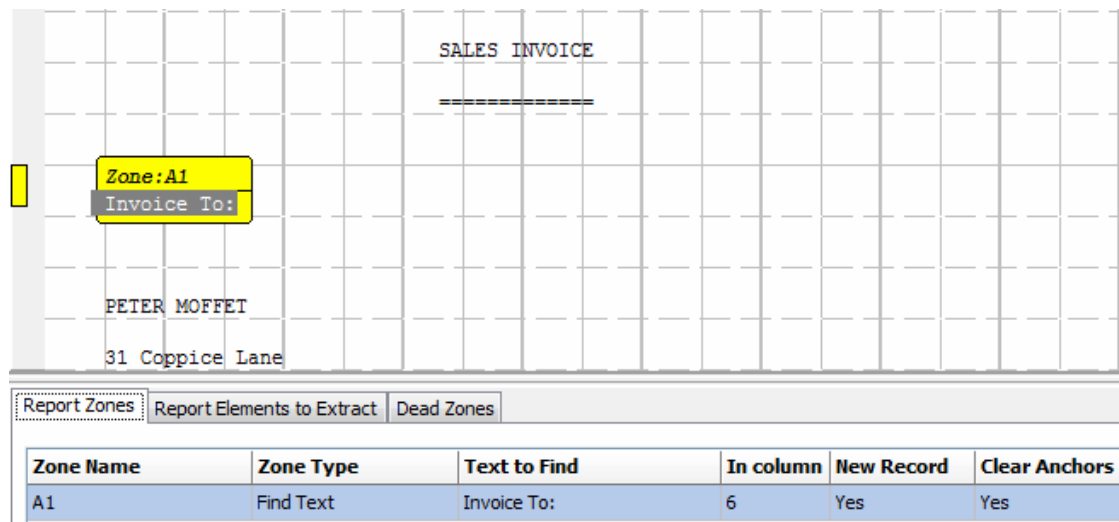
If this is set, any other anchors will be cleared once this anchor has been found. This should be set except for rare circumstances.

## Changing an Anchor

---

Once an anchor has been added, the Report Zone Editor displays the anchor in a yellow box. You can edit the anchor settings by double clicking the box.

The anchors are also listed in the Report Zones tab beneath the report panel. You can edit or delete anchors from this list by using the Zone menu.



Zone Name	Zone Type	Text to Find	In column	New Record	Clear Anchors
A1	Find Text	Invoice To:	6	Yes	Yes

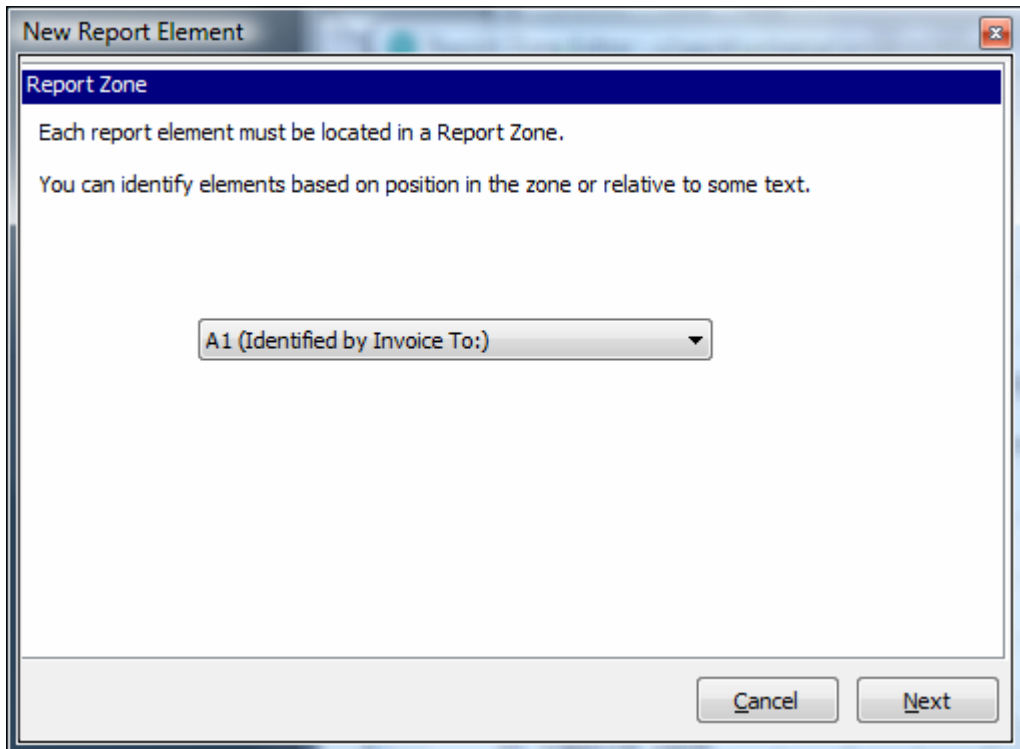
## Adding a Simple Report Element

---

The Report Zones determine when and where mvPDF will start to scan for data to extract. The Report Elements specify exactly which data needs to be extracted from the zone, and where it should be placed.

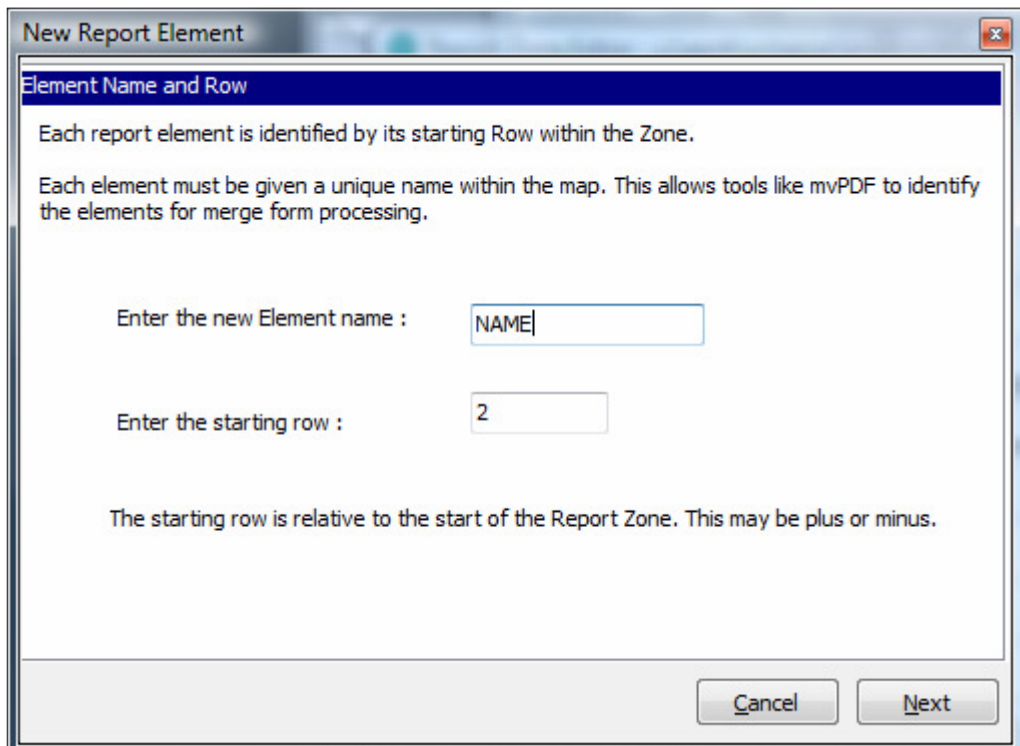
To add a new Report Element to your map, you can select the Add option from the Element menu or click the Add Element button. As with the Report Zone, it is very much easier if you select the text that makes up an element in the report panel first. If the element is multiple line, you only need select the first line of the element.

This presents the Add Report Element wizard to walk you through the process of defining an element:



**Report Zone:**

The first page of the wizard asks which report zone should contain your element. From here, any offset calculations are presented with respect to that zone.



### Element Name:

Each extraction element must be given a unique name. If you are passing this directly into an mvPDF Merge Form, the element name will match the name of an object in the merge form.

### Starting Row:

The starting row determines where mvPDF will begin to extract this element. As mvPDF parses the report line by line, it keeps a list of found anchors and checks current row against the elements attached to those zones to see which ones to extract. The starting row is therefore an offset from the Anchor row. If you need to reference some text that appears before the anchor, for example where a name and address appear above a reference, you can specify a negative offset.

**New Report Element**

**Element Starting Column**

Each report element is identified by a starting column for the text.  
You can specify an exact column number (starting from 1) or identify the starting point relative to some text such as a label.

How do you want to identify the starting column :

by absolute column position  
 relative to some text

Enter the column number or text to locate :

6

Cancel Next

### Starting Column:

The starting column defines where on the current line, starting from column 1, mvPDF should start to extract the element.

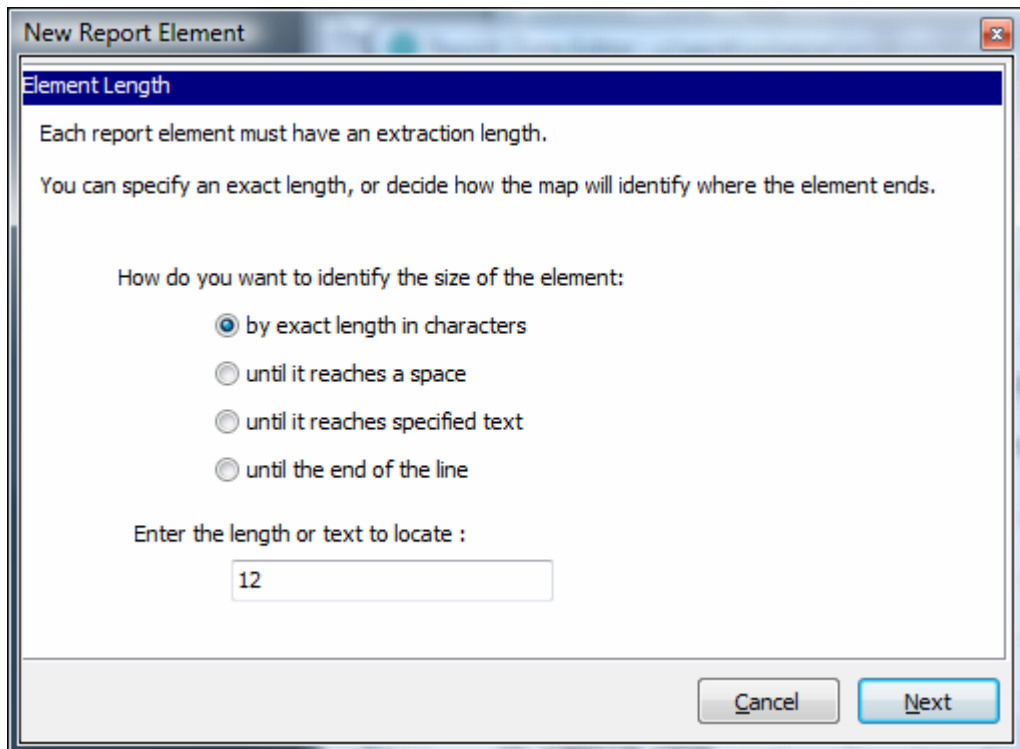
This can be an absolute position or a position relative to a Zone anchor or to a label.

For an absolute column position, enter the column number in the box below.  
For an anchor column offset, specify the anchor name with a plus or minus offset (A1+10).

Alternately, if the value follows a label you can specify the label and mvPDF will extract the characters following on the line. For example, if the report contained a line:

Reference : 01233445

You could set the starting column type to 'relative to some text' and the text to 'Reference : '

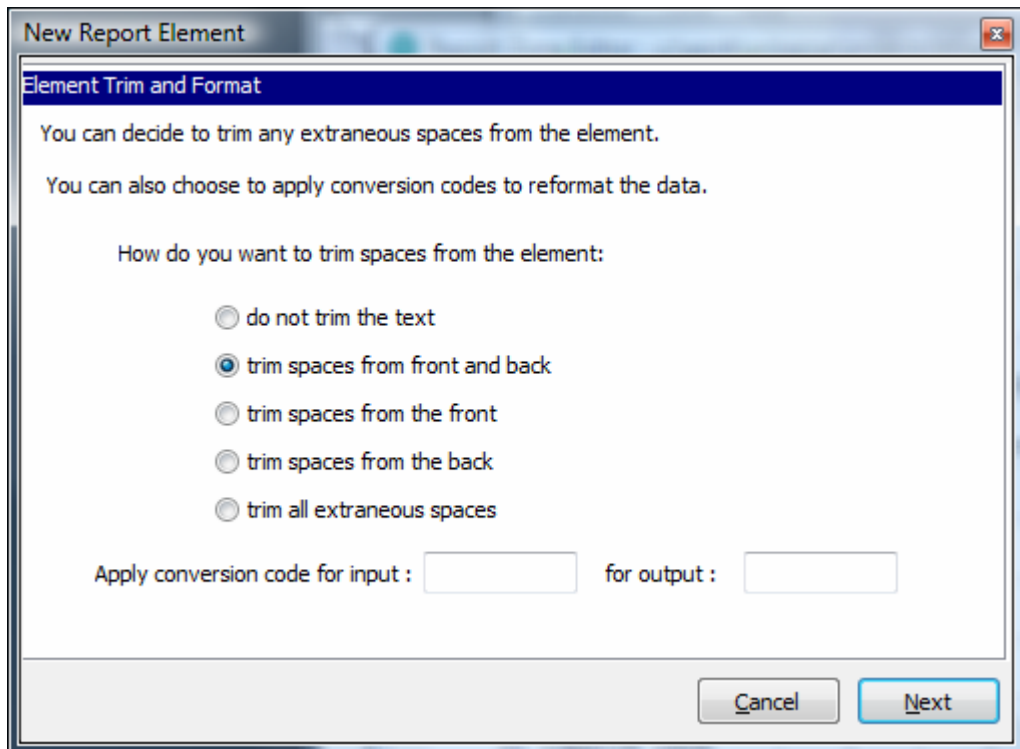


#### Element Length:

You have a similar set of choices to determine the length of the element that mvPDF should extract, starting from the column set in the previous page.

The data may be a fixed length (including spaces), may be a single figure or code that will be terminated by a space, or may continue to the end of the line. You can also specify text that marks the end of that column, for example if you have tabular data separated by vertical bars (|) or several labels on the same line, or if your value is embedded into some text:

Please quote reference number **012345** and ask for **Dawn** in customer support.

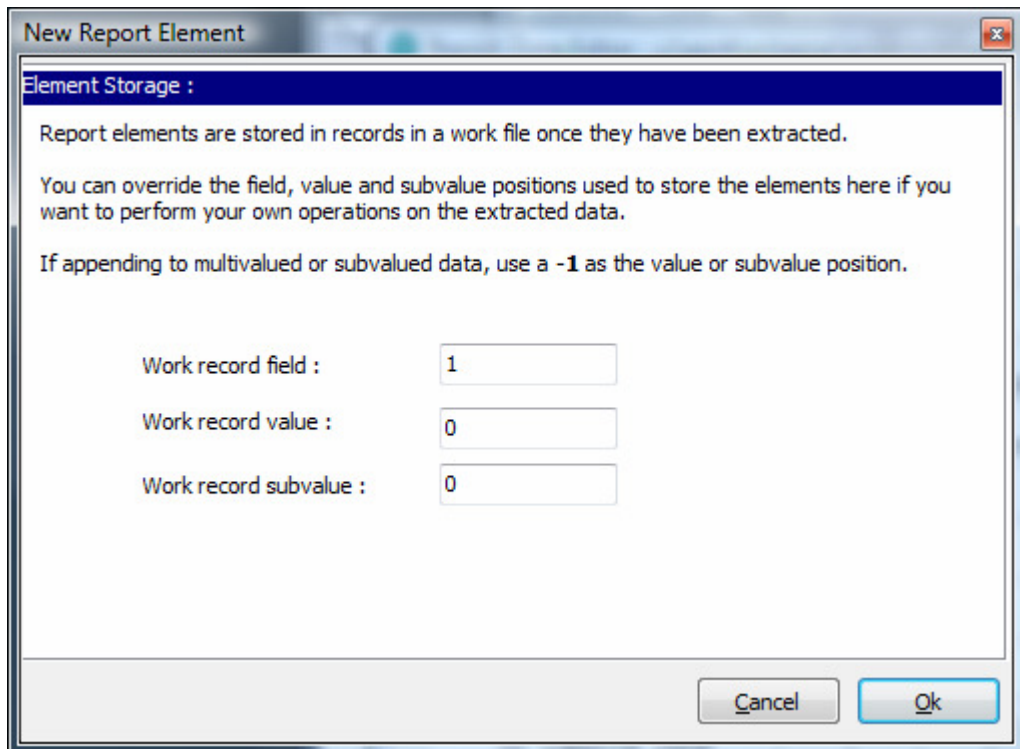


**Element Trim:**

If you specify an element length that is fixed you may need to trim the extracted data to remove any leading or trailing spaces. mvPDF offers a range of options for trimming the start, end or all extraneous spaces.

**Conversion Codes:**

If you are processing the data forward after extraction, you may wish to apply either or both of an input and output conversion to the data. mvPDF will apply an input conversion first, for example to allow you to change a date format by applying a generic date (D) conversion on input and your own date format conversion back out.



**Storage:**

When mvPDF extracts the data from a report, it writes that data to one or more records in a work file. You can tell mvPDF how the record should be updated, by specifying the field, value and subvalue numbers that should be updated for this element.

Once the Add New Element wizard has been completed, the Report Zone Editor displays the element in blue. You can edit the element details by double clicking on the element in the report panel, or from the Report Elements tab below.

Name	Column Type	Column Value	Zone	Row Value	Extract Type	Extract Value	Trim Style
NAME	Column Position	6	A1	2	Base on length	12	Both ends

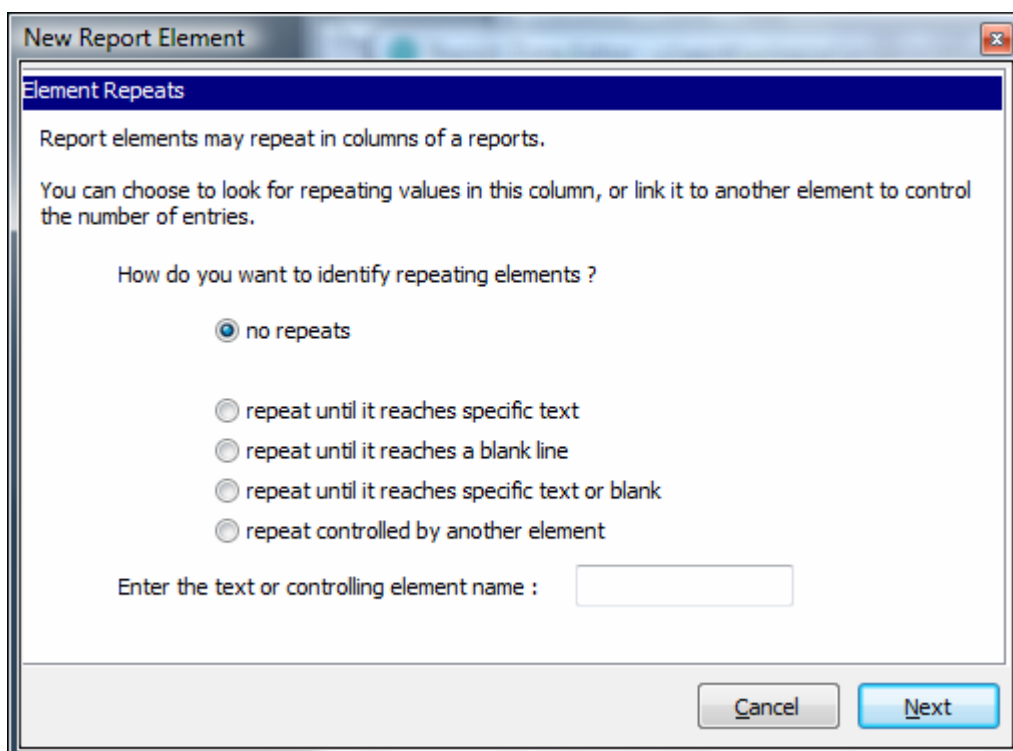
## Adding Repeating Elements

---

mvPDF has special handling for repeating elements in reports, such as the detail lines on an invoice or the columns in an enquiry listing.

mvPDF only handles repeating fields that are represented in columnar format, with repeating values occupying successive rows. To determine how elements can repeat, mvPDF needs to know when the last line in the repeating series has been reached.

Repeating Fields are specified in the Add Report Element Wizard.



The screenshot shows a dialog box titled "New Report Element" with a tab labeled "Element Repeats". The dialog contains the following text and options:

Report elements may repeat in columns of a reports.

You can choose to look for repeating values in this column, or link it to another element to control the number of entries.

How do you want to identify repeating elements ?

- no repeats
- repeat until it reaches specific text
- repeat until it reaches a blank line
- repeat until it reaches specific text or blank
- repeat controlled by another element

Enter the text or controlling element name :

At the bottom right, there are two buttons: "Cancel" and "Next".

### Element Repeats

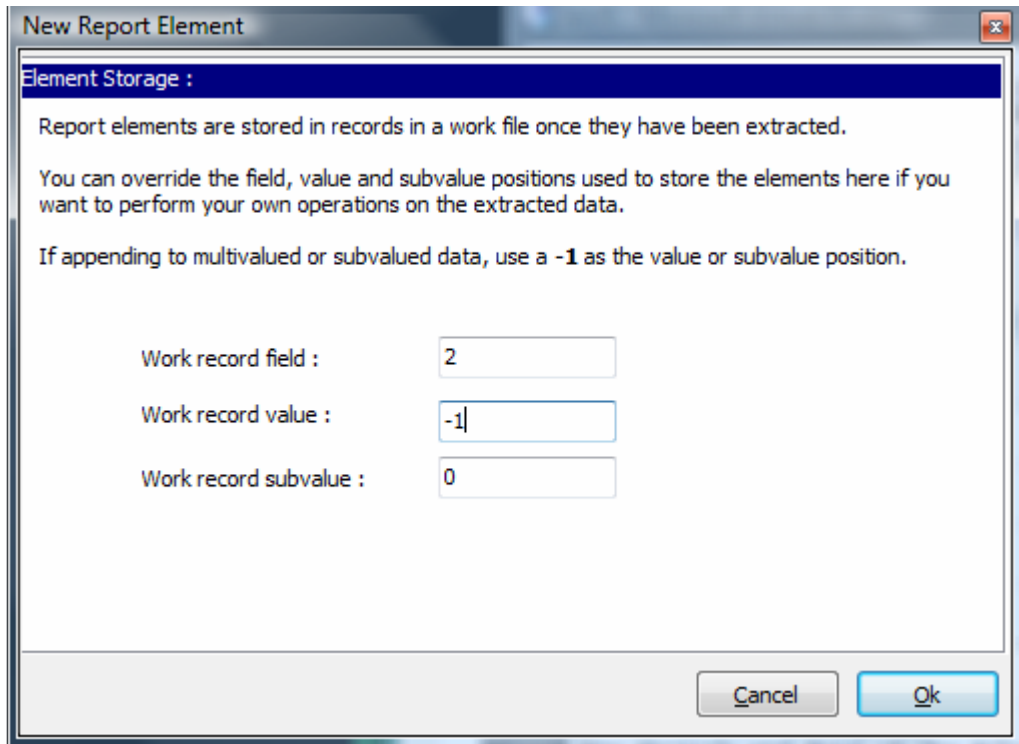
You can specify a fixed number of lines over which the column will repeat, or the column may repeat until it reaches a blank line or a marker such as a total bar.

If you have a set of elements that repeat as one unit, as with the detail lines on an invoice, it is better to link these. This prevents any possibility that one element may be truncated early if it has a blank line accompanying a value in another element.

To link elements together, you specify a 'repeat controlled by another element' in the dependent column and set the name of the controlling element in the text box below. It is important to select a controller that will not include blank values in the series, such as a product code for an invoice line.

## Repeating Element Storage

If you have repeating elements, there are two choices for how you can store these in the work record created.

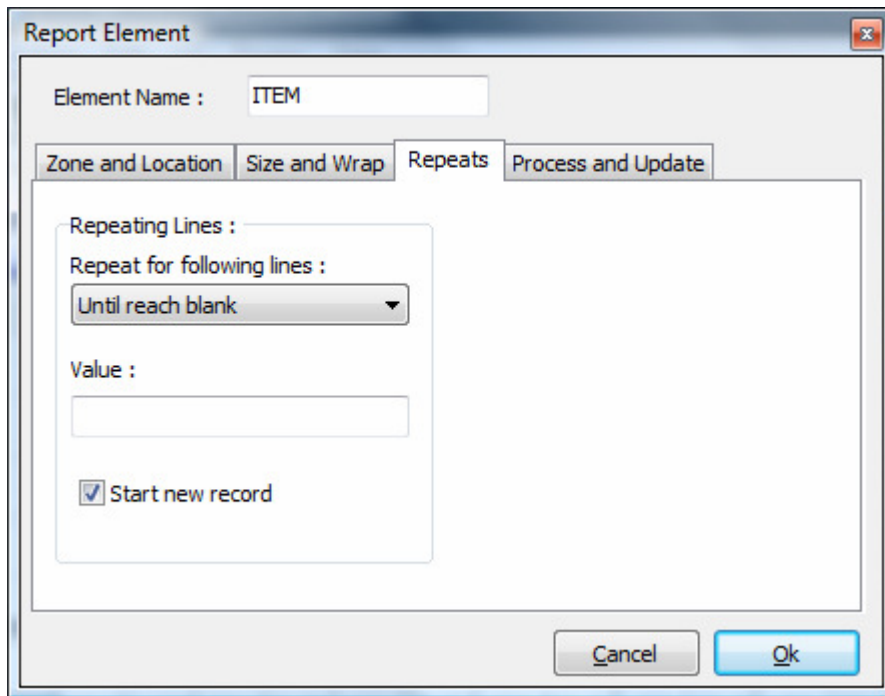


The screenshot shows a dialog box titled "New Report Element" with a close button in the top right corner. The dialog has a blue header bar with the text "Element Storage :". Below the header, there is instructional text: "Report elements are stored in records in a work file once they have been extracted. You can override the field, value and subvalue positions used to store the elements here if you want to perform your own operations on the extracted data. If appending to multivalued or subvalued data, use a -1 as the value or subvalue position." Below this text are three input fields: "Work record field :" with the value "2", "Work record value :" with the value "-1", and "Work record subvalue :" with the value "0". At the bottom right of the dialog are "Cancel" and "Ok" buttons.

You can create multivalued or sub-multivalued fields, by setting the Value Number or SubValue number to -1. This will append elements to the work record field or value. If the repeats are linked, mvPDF will use the controller value or subvalue to align the data. This is usually the best option if you handling subordinate details, such as the detail lines on an invoice.

If you are processing a columnar report or wish otherwise to separate out each repeat, you can set the Create New Record option in the element details. This will create a new Work Record for each repeat, and you should not set the value or subvalue to -1.

The Create new Record option is not included in the Add New Element wizard. It is only available when you edit an existing element:



## Repeats and Text Wrapping

An added complication arises from the fact that some fields may wrap onto following lines. Text wrapping needs to be handled separately to repeating, as in the example below:

```

SORT BOOK_TITLES 08:18:31pm 03 Jun 2009 PAGE 1
Key.. Title..... Author Name.....

  1 Just William: No. 6 (BBC Radio RICHMAL CROMPTON
    Collection)
  2 Just So Stories RUDYARD KIPLING
  3 The Duchess of Malfi (abridged JOHN WEBSTER
    version)
  4 Great Expectations CHARLES DICKENS
  5 The Importance of Being OSCAR WILDE

```

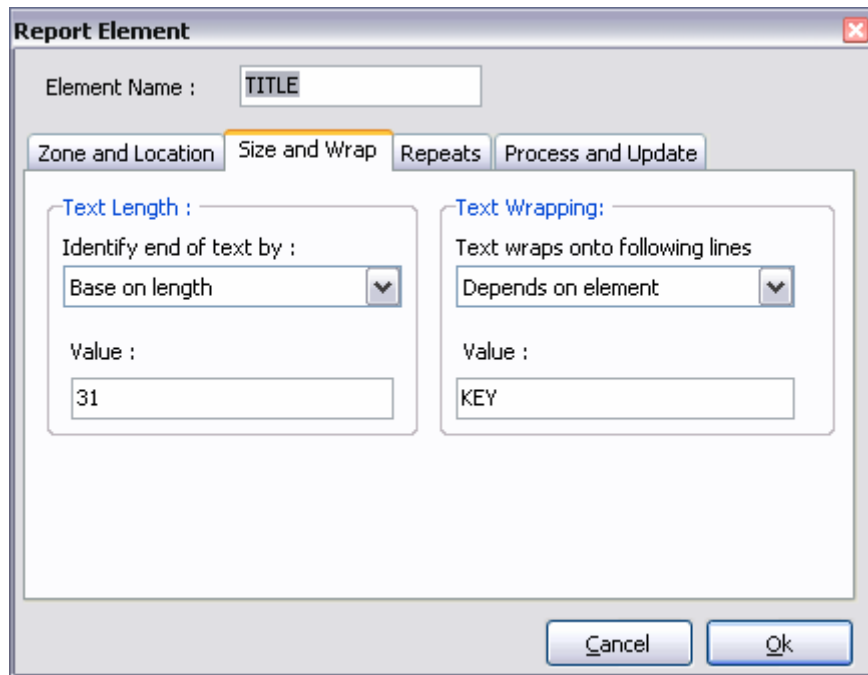
This listing above displays the two forms of continuations supported by mvPDF: text wrapping and repeating elements.

The Title column spills over two lines for the first and third entries 'Collection)' and 'version)'. mvPDF should treat these as continuations of the lines above. This is handled by the text wrapping specification.

The Key, Title and Author all repeat onto following lines, with the breaks caused by the text wrapping, until they reach a blank line at the end of the page or of the series. This is handled by the repeat specification.

To control this, you must tell mvPDF that the Title column is subject to text wrapping. However, mvPDF cannot know where the text wrapping for each line ends, just by looking at the Title column. The only way to identify that the wrapped line holding the 'Collection)' belongs to the title above and should not be treated as a separate repeat, is by looking at the adjacent Key column.

To add text wrapping to a report map, you must tell mvPDF which element defines where the wrapping will end.



For a series of repeats such as the one above, mvPDF will append the following lines so long as the controlling element does not contain a value.

## Adding Dead Zones

---

Dead Zones are sections of a report page that have no meaningful data. Dead Zones are normally reserved for page headers and footers.

If you are processing a series of repeats that can continue over more than one page, the blank lines of the page footer and the text of a page header would interrupt the series and could force mvPDF to prematurely finish processing the repeats if they are set to finish on a blank line.

To prevent this, you can define a Dead Zone for the start and/or the end of a page. Dead Zones are skipped by mvPDF when parsing through a report.

To add a Dead Zone to the report map, select the Add option from the Dead Zone menu.

The image shows a dialog box titled "Dead Zone". It contains the following fields and options:

- Starting Line from Top of Page: 1
- Number of Rows to Disacrd: 6
- Treat first page the same

Buttons: Cancel, Ok

Dead Zones are identified by absolute row number from the top of the page. Since they are skipped without processing, it is not possible to dynamically determine different dead zones on different pages.

## Saving a Report Map

---

You can save your report map at any time as you are working on your report. To save the map you should select the Save or Save As options from the File menu. Each report map must be given a unique name.

The maps can be reused if you have several reports with the same layout, and more than one map can be applied to the same report, so there is no link recorded between the report and the report map. It is recommended that you call the map by a meaningful name that will allow you to remember which map accompanies which report. You should also avoid using spaces in the Report Map name.

The map is saved to the PDFMAPS file on the server.

## Loading a Report Map

---

You can load an existing report map by selecting the Open Map option from the File menu.

This presents a list of the existing maps for you to select.

## Running the Report Mining

---

Once you have defined your maps, you can use this with the spooler capture commands to rip data to a work file or to feed directly into an mvPDF Merge Form.

## Report Mining to a Work File

---

The SET.PDF command signals the start of the report capture for mvPDF.

To run a report through the mvPDF Report Mining facility, run the SET.PDF command specifying the name of the map and optionally the name of a work file thus:

**SET.PDF /MAP mapName /TARGET workfileName**

For example:

SET.PDF /MAP INVOICE /TARGET INVOICEDATA

If you do not specify a work file name, mvPDF will use the generic workfile PDFWORK.

mvPDF will process the report when it encounters a RESET.PDF command:

**RESET.PDF**

To prevent different sets of report data from overwriting each other in the same workfile, mvPDF writes the work file records with a key in the format:

job\_number : "." : sequence\_number

If you wish to apply your own processing following the RESET.PDF command, for example to generate your own exports and formatting of the extracted data, you can find your last job number by calling the pdfJobNumber subroutine:

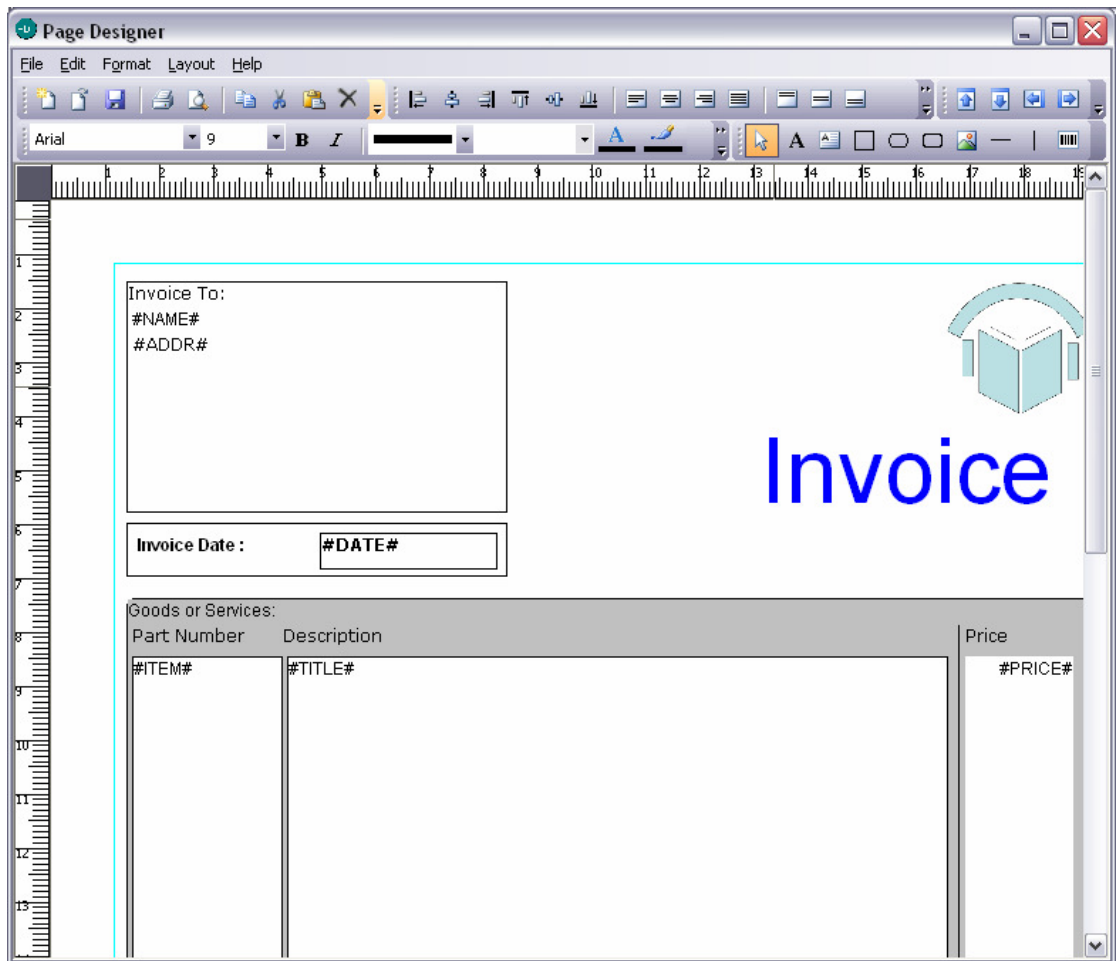
Call pdfJobNumber(lastJobNumber)

## Report Mining to a Merge Form

---

mvPDF Merge Forms provide a convenient way to generate new mvPDF reports and forms using the page designer.

Merge Forms are created using the same Page Designer used for Watermark pages, but can include objects whose content is filled in at run time. These objects each specify the name of a merge field enclosed in hashes, as below:



The mvPDF Report Mining facility can direct the results of a report extraction to a merge form, using the names of the elements in the Report Map to identify the merge fields in the form.

For this you must specify an additional argument to the SET.PDF command that provides the name of the merge form.

#### **SET.PDF /MAP pathToMapFile**

If the merge has continuation pages, you can specify that as a second merge form. You will also need to complete the necessary arguments that determine the name and settings for the document produced or the name of a destination printer:

```
SET.PDF /MAP INVOICE /MERGE "c:\forms\invoice.xml" /MERGE "c:\forms\contine.xml"  
/NAME "invoice.pdf" /PATH "c:\invoices"
```

You can read more about Merge Forms and merge processing in the mvPDF User Guide.