White Paper

// Regular Expressions used in uniFLOW and HelixPW

Version 1.1

23-Aug-2012

Version: All

Version: All

Version: All

Versioning

Versioning	Version	Author(s)	Revision date	Reviewer(s)
	1.0	Klaus Langenberg	7-APR-2012	Thomas Lemmer
	1.1	Felix Schlick	23-AUG-2012	Thomas Lemmer
WP Name	Regular Expressions used in uniFLOW and HelixPW			
Knowledge Base	MOMKB-635 https://web.nt-ware.net/its/browse/MOMKB-635			
File Name	WP - Regular Expressions used in uniFLOW and HelixPW - V1.0.PDF			
Technologies concerned	uniFLOW - all versions			
Short Summary	This White Paper lists the set of Regular Expressions that are used in uniFLOW and HelixPW.			
Document changes	**Version**	**Changes**		
	1.1	Added footnote concerning escaping of "\" and "/"		

Disclaimer

Company and product names mentioned herein are registered or unregistered trademarks of their respective companies. Mention of third-party products is for information purposes only and constitutes neither an endorsement nor a recommendation. NT-ware assumes no responsibility with regard to the performance or use of these products. Also, NT-ware makes no claim to these trademarks. Any use of trademarks, logo, service marks, trade names, and product names is prohibited without the written permission of the respective owners.

Adlib Software of Adlib Software; Adobe®, Adobe® Reader, Acrobat®, Distiller®, PostScript® and products of the CREATIVE SUITE(S) of Adobe Systems Incorporated; Apple®, the Apple® logo, Mac®, Mac OS®, Macintosh®, iPhone®, iPad® and AirPrint® of Apple Inc.; CANON, imageRUNNER, imageRUNNER ADVANCE, MEAP, CPCA, AMS, iW AMS, iW Desktop, iSend, iW SAM of Canon Inc.; Crystal Reports of Business Objects SA, as of July 1, 2008: BusinessObjects of SAP; eCopy™, eCopy ShareScan®, and eCopy ScanStation® of Nuance Communications, Inc.; Foxit Reader of Foxit Corporation; Google Docs of Google Inc.; Helix™ Production Workflow is a trademark of NT-ware Systemprogrammierung GmbH; Hewlett Packard, HP, LaserJet, and PCL of Hewlett-Packard Company; iOS® of Cisco Technology Inc.; I.R.I.S. Group s.a.; JAWS pdf courier™ are trademarks of Global Graphics SA.; Microsoft®, Windows®, Windows Vista®, Windows 7®, Internet Explorer®, Internet Information Server, Microsoft® Word, Microsoft® Excel, SQL Server® of Microsoft Corporation; Neevia Document Converter ProTM of Neevia Technology; NetWare, Novell®, Novell eDirectory® of Novell Inc.; OpenOffice.org™ of Oracle Corporation; PAS™ of Equitrac Corporation; PosterJet of Eisfeld Datentechnik GmbH & Co. KG; Red Titan EscapeE of Red Titan Limited; NETHAPHOR ®, SiteAudit™ are trademarks of NETAPHOR SOFTWARE Inc.; Therefore™ of Therefore; UNIX® of The Open Group; uniFLOW OM®, uniFLOW™, uniFLOW Serverless Secure Printing®, MIND®, microMIND®, and MiCard® are registered trademarks of NT-ware Systemprogrammierung GmbH; pcProx®, AIR ID® are registered trademarks of RFIdeas Inc.Readers; CASI-RUSCO® is registered trademark of ID Card Group; Radio Key® is registered trademark of Secura Key; GProx™ II is unregistered trademark of Guardall; HID® ProxHID is registered trademark of HID Global Corporation; Indala® is registered trademark of Motorola; ioProx™ is unregistered trademark of Kantech.

All other trademarks, trade names, product names, service marks are property of their respective owners and are hereby acknowledged.

While every precaution has been taken in the preparation of this document, NT-ware assumes no responsibility for errors or omissions, or for damages resulting from the use of information contained in this document or from the use of programs and source code that may accompany it. NT-ware does not assume any responsibility or liability for any malfunctions or loss of data caused by the combination of at least of one NT-ware product and the used operation system and/or third-party products. In no event shall NT-ware be liable for any loss of profit or any other commercial damage caused or alleged to have been caused directly or indirectly by this document.

In addition, this manual provides links to the sites of affiliated or independent companies and certain other businesses. NT-ware is not responsible for examining or evaluating, and NT-ware does not warrant the offerings of, any of these businesses or individuals or the content of their websites. NT-ware does not assume any responsibility or liability for the actions, product, and content of all these and any other third parties. You should carefully review their privacy statements and other conditions of use.

PLEASE NOTE: Serious problems might occur if you modify the registry of your Windows operating system incorrectly. These problems might require that you reinstall the operating system. We strongly recommend to always back up the registry of your Windows operating system before applying changes to it, just in case you do something wrong. NT-ware does not assume any responsibility or liability for any impact on the operating system after changing the Registry. You understand and accept that you use this information and modify the registry of your Windows operating system at your own risk.

Donnerstag, 23. August 2012, Bad Iburg (Germany)

Symbols

Text Styles

Text that appears in this style is used for screen text that appears in the uniFLOW user interface and on user interface controls.

Text that appears in this style is used for User entries on screen, text that the user actually has to type in.

Text that appears in this style is used for hyperlinks to an external web page, or internal links to other pages of this manual.

```
Text that appears in this style is used for code examples: XML code, variables or regular expressions.
```

Pictograms

Important note: Information that is crucial for the correct functioning of the uniFLOW software.

External manual: Pointer to additional manuals for third party hardware or third party software.

Region Specific Feature: In case some features of uniFLOW are not universally available, this icon will indicate it.

Link to an external reference within the WWW.

Detailed explanation of configuration settings or operational procedures.

Copyright and Contact

In case of errors or improvement suggestions please contact documentation@nt-ware.com.

Contents

1 Regular Expressions 2

1 Regular Expressions

The passed regular expression consists of meta characters and regular characters. Regular characters are matched directly. Meta characters have a meaning according to the following table.

Meta-character	Meaning	Example
`.`	Match any single character.	
`[ ]`	Defines a character class. Matches any character inside the brackets.	`[abc]` matches "a", "b", and "c"
^	If this metacharacter occurs at the start of a character class, it negates the character class. A negated character class matches any character except those inside the brackets.	`[^abc]` matches all characters except "a", "b", and "c").
	If ^ is at the beginning of the regular expression, it matches the beginning of the input.	`^[abc]` will only match input that begins with "a", "b", or "c".
-	In a character class, indicates a range of characters.	`[0-9]` matches any of the digits "0" through "9".
?	Indicates that the preceding expression is optional: it matches once or not at all.	`[0-9][0-9]?` matches "2" and "12".
+	Indicates that the preceding expression matches one or more times.	`[0-9]+` matches "1", "13", "666", and so on.
*	Indicates that the preceding expression matches zero or more times.	

`??, +?, *?`	Non-greedy versions of ?, +, and *. These match as little as possible, unlike the greedy versions which match as much as possible.	given the input "<abc><def>",
		`<.*?>`
		matches "<abc>" ,
		`<.*>`
		matches "<abc><def>".

`( )`	Grouping operator.	`(\d+,)*\d+` matches a list of numbers separated by commas (such as "1" or "1,23,456").
`{ }`	Indicates a match group.	

Meta-character	Meaning	Example
\	Escape character: interpret the next character as a literal. Also used for abbreviations (such as \a for any alphanumeric character; see further below). *For escaping the back slash character, see footnote.	[0-9]+ matches one or more digits, but [0-9]\+ matches a digit followed by a plus character.
	If \ is followed by a number n, it matches the nth match group (starting from 0).	<{.*?}>.*?</\0> matches "<head>Contents</head>".
$	At the end of a regular expression, this character matches the end of the input.	[0-9]$ matches a digit at the end of the input.
|	Alternation operator: separates two expressions, exactly one of which matches.	T|the matches "The" or "the"
!	Negation operator: the expression following ! does not match the input.	a!b matches "a" not followed by "b".
\a	Any alphanumeric character.	Shortcut for ([a-zA-Z0-9])
\b	White space (blank).	Shortcut for ([\t])
\c	Any alphabetic character.	Shortcut for ([a-zA-Z])
\d	Any decimal digit.	Shortcut for ([0-9])
\h	Any hexa-decimal digit.	Shortcut for ([0-9a-fA-F])
\n	Newline.	Shortcut for (\r|(\r?\n))
\q	A quoted string.	Shortcut for (\"[^\"]*\")|(\'[^\']*\'))
\w	A simple word.	Shortcut for ([a-zA-Z]+)
\z	An unsigned integer.	Shortcut for ([0-9]+)

* You can use backslash and forward slash as literal by escaping them themselves:

<#bs#> = <\>

<#fs#> = </>

This is particularly useful when using the device agent Mind SMTP Email Control in conjunction with a Cycos MRS server, where addresses can contain both kinds of slashes.