
PREMIER TRANSFORMATION FUNCTIONS TOOLBOX

USER / DEVELOPER GUIDE

8.6.1 Release

Version 2010.05.31.01

ABOUT FYGHTSOFT

FYGHTSOFT PRODUCTS

Our products are created out of real world necessity and based on best business practices. Each of our products aims to solve a business or integration problem that existed at one or more of our current or previous clients. Our products are field tested and battle ready. In order to keep our solutions fresh and products relevant we continually ask our clients and potential buyers of our products to tell us about their integration issues. If you have an idea for a product or an existing integration issue, let us know. Our products are designed to help.

COPYRIGHT

All rights and ownership of FyghtSoft products are reserved by FyghtSoft, LLC.

LIMIT OF LIABILITY / DISCLAIMER

While every precaution has been taken in the preparation of this article, the publisher and author assume no responsibility for errors or omissions, or for damages resulting from the use of the information contained herein.

TRADEMARKS

Designations used by companies to distinguish their products are often claimed as trademarks or registered trademarks. In all instances in which FyghtSoft is aware of such claim, the product names appear in initial capital or all capital letters. Readers, however, should contact the appropriate companies for more complete information regarding trademarks and registration.

WEBSITE

<http://www.fyghtsoft.com>

SALES & SUPPORT

If you have any questions regarding the purchase of this product or require additional product licenses, please contact our sales office, Monday-Friday from 9am – 4pm EST at 704-904-3667 or sales@fyghtsoft.com.

If you have technical issues or need to report a problem with the software please log the issue on our support forum at <http://www.fyghtsoft.com/forums/>.

PARTNERS

FyghtSoft maintains relationships with several organizations, however, mention of any organization within this document does not construe entitlement, relationship, partnership, ownership, or management of any entity unless otherwise explicitly stated.

Table of Contents

About FyghtSoft	2
Getting Started with Development	4
Introduction	4
Overview.....	4
License Granted.....	4
Certified Environments Matrix.....	4
Pre-Requisites	4
Function Breakdown	6
Overview.....	6
References	9
Conclusion.....	9
Demo Quickstart	10
Individual Function Operations.....	10
Random Functions Caveats.....	10
URLEncodeDecode	10
Is_ValidEmailAddress.....	10
ChecksumCRC	11
Conclusion.....	11

GETTING STARTED WITH DEVELOPMENT

INTRODUCTION

The FyghtSoft Premier Function Toolbox for Informatica PowerCenter integrates seamlessly into the Informatica Designer client tool. You will take advantage of a familiar set-up, now you just have a bit more power as a developer.

OVERVIEW

The Informatica Premier Transformation Functions Toolbox provides a well needed extension to the Transformation Expressions function list currently included with Informatica PowerCenter. Usage of these functions is limited only to the creative solution imagination of the end users; however a sample mapping and data source are provided to test the functions immediately upon your installation of the product.

LICENSE GRANTED

Purchase of this product entitles your organization to a single (1) server production license. If you have more than one Informatica Server you will need to purchase additional products/licenses to satisfy each server instance. Please contact our Sales department as volume pricing is available. If you have any questions regarding your license or any compliance thereto, please contact our sales department.

CERTIFIED ENVIRONMENTS MATRIX

In order to comply with our standards as well as those of Informatica, our software has been tested thoroughly and rigorously.

Currently this product is certified (tested and proven) on the following environments:

	PowerCenter 8.6.x	PowerCenter 9.x	Others Coming Soon
Windows Server 2003	X	X	
Windows Server 2008	X	X	
Linux			
Sun Solaris			

PRE-REQUISITES

As a pre-cursor to reading and utilizing this guide, one should have read and completed the installation of the FyghtSoft Premier Function Toolbox into the Informatica server and client environments. Please reference the Installation Guide included with this product download for this information.

In addition, as a developer, one will need developer level access to Informatica PowerCenter Designer and the repository under which the Premier Function Toolbox plug-ins were installed. Please confer with the Informatica administrator if necessary to attain this information.

FUNCTION BREAKDOWN

OVERVIEW

The Informatica PowerCenter Premier Functions Toolbox extends the existing transformation expression functions in a myriad of ways. Some of these are minor while others provide major functionality that reduces development time and thus cost. The Premier Functions Toolbox also eliminates the need for development of code outside of the Informatica development environment to be performed. Below is a summary breakdown of each function included in the Premier Functions Toolbox. Please see the accompanying demo included with the product download and/or the video demo found in the resources section of our website.

Function Name		
ChecksumCRC	Arguments	1 to 50 (50 max)
		ChecksumCRC(i_key, i_customerKey, i_mgrKey, i_saleAmount, ...)
	Returns	String (Max 100 characters)
	Description	<p>All input arguments are passed through the function and used to generate the checksum which is returned as a single scalar result string, CheckSumCRC.</p> <p>A cyclic redundancy check (CRC) or polynomial code checksum is a non-secure hash function designed to detect accidental changes to raw computer data, and is commonly used in digital networks and storage devices such as hard disk drives. It has also become widely adopted in quickly accessing changes in Data Warehousing Slowly Changing Dimension and Change Data Capture.</p> <p>Other MD5 hash algorithms used to create a checksum against data table row columns have been used in the past. However, our CheckSumCRC function is superior in consistency and provides a more open implementation.</p>
GUIDGenerator	Arguments	None
		GUIDGenerator()
	Returns	A 36-character string (xxxxxxxx-xxxx-xxxx-xxxx-xxxxxxxxxxxx)
	Description	<p>Provides a randomly computed Globally Unique Identifier based on a superior algorithm complying with the Universally Unique Identifier (UUID) standard.</p> <p>A globally unique identifier or GUID is a special type of identifier used in software applications to provide a unique reference number. The value is represented as a 32 character hexadecimal character string, such as {21EC2020-3AEA-1069-A2DD-08002B30309D} and usually stored as a 128 bit integer.</p>

		<p>The primary purpose of the GUID is to have a totally unique number. Ideally, a GUID will never be generated twice by any computer or group of computers in existence. The total number of unique keys (2^{128} or 3.4×10^{38}) is so large that the probability of the same number being generated twice is extremely small, and certain techniques have been developed to help ensure that numbers are not duplicated.</p>
Is_LeapYear	Arguments	Arg1 as Integer
	Returns	Integer (0 = false, 1 = true)
	Description	Provides a mechanism to determine if a year falls into the elite group of years known as “Leap” years.
Is_Palindrome	Arguments	Arg1 as String
		Is_Palindrome('najan')
	Returns	
	Description	<p>This function takes into account a string’s composition and determines if it is canonically recognized as a Palindrome.</p> <p>A palindrome is a word, phrase, number or other sequence of units that can be read the same way in either direction (the adjustment of punctuation and spaces between words is generally permitted). However, this function only takes into account single words and will fail any input with spaces.</p> <p>Uses of this function include mass scanning for unique website names such as Kayak (http://www.kayak.com), etc. due to their unique mnemonic properties.</p>
Is_ValidEmailAddress	Arguments	Arg1 as String
		Is_ValidEmailAddress('steve.jobs@apple.com')
	Returns	Integer (0=false, 1=true)
	Description	<p>Complies with all current email standards specifically those identified in detail by The Internet Engineering Task Force, RFC2821 and RFC2822 (see references below). This function uses a strong parsing algorithm to assess the format of an email address parsing its structure.</p> <p>Note: This function does not attempt to connect to the email server on which the email address argument resides to verify the email address is live, disabled, or otherwise.</p>
RandomString (Password Generator)	Arguments	<p>Length as Integer, IncludeNumbers as Integer ([0 1]), IncludeSpecialCharacters as Integer ([0 1])</p> <p><i>Max Length: 100</i></p>

		RandomString(9, 1, 0) <i>Example Output: LE7aq9Pvm</i>
	Returns	String (Max Length 100)
	Description	<p>The RandomString function, also known as the Password Generator, provides an unparalleled capability in Informatica PowerCenter. This function returns a randomly generated string varying in length and character composition based on the static or dynamic arguments it is passed.</p> <p>The string generator is based on three character sets: Alpha, Numeric, and Special Character. Alpha is the base set and provides by default randomization of upper and lower case letters (A-Za-z). The two remaining character sets are selectable by passing in the appropriate flag (as integer [0 1]) into the last two arguments respectively.</p>
StringContains	Arguments	BaseText as String, SearchText as String
		StringContains('This is our FyghtSoft product', 'our')
	Returns	Integer (0=false, 1=true)
	Description	This function is similar to the Microsoft .NET String function for testing if a base string contains a search string. If the search string is found within the base string then the function returns a value representing a true (1) evaluation, otherwise false (0).
URLEncodeDecode	Arguments	Arg1 as String, Arg2 as String
		URLEncodeDecode(Arg1 as String, ['ENCODE' 'DECODE'] as String)
	Returns	String (Max 500 characters)
	Description	<p>URL encoding converts characters into a format that can be safely transmitted over the Internet. URLs can/should only be sent over the Internet using the ASCII character-set.</p> <p>Since URLs often contain characters outside the ASCII set, the URL string has to be converted. URL encoding converts the URL into a valid ASCII format.</p> <p>URL encoding replaces unsafe ASCII characters with "%" followed by two hexadecimal digits corresponding to the character values in the ISO-8859-1 character-set.</p> <p>URLs cannot contain spaces. URL encoding normally replaces a space with a + sign. This function uses a + sign and not %20 for spaces.</p> <p>The decoding option is available in the function by passing "DECODE" as the second argument instead of "ENCODE". It reverses the encoding.</p>
WordCount	Arguments	Arg1 as String
		WordCount('No one can tell you what the Matrix is.')
	Returns	Integer

	Description	<p>A very simple but robust function that should be in every Informatica developer's toolbox. This function takes the simple "length" function and turns it up a notch. Now analysis and expressions can be made smarter to identify column meta data that is too long or too short based on the count of words within a string not just rudimentary length of the string.</p> <p>Counting of words is based solely on spaces within the string/text being assessed by the function.</p>
--	-------------	--

REFERENCES

For this and more information regarding some technical definitions of functionality provided within the Premier Function Toolbox, please reference the following:

- URLEncodeDecode: http://www.w3schools.com/tags/ref_urlencode.asp
- CheckSumCRC: http://en.wikipedia.org/wiki/Cyclic_redundancy_check
- Palindrome: <http://en.wikipedia.org/wiki/Palindrome>
- Is_ValidEmailAddress:
 - <http://tools.ietf.org/html/rfc2821>
 - <http://tools.ietf.org/html/rfc2822>

CONCLUSION

The above breakdown of each function provides an immediate syntactical reference and brief understanding for the overall capabilities of this product. There are very few caveats inherent within the custom functions provided within the toolbox. These caveats are discussed in a section found towards the end of this document. Importing and running the demo included with this product is highly recommended in order to see how the functions operate. Please also reference the quickstart videos and guides from <http://www.fyghtsoft.com/index.php/resources>.

DEMO QUICKSTART

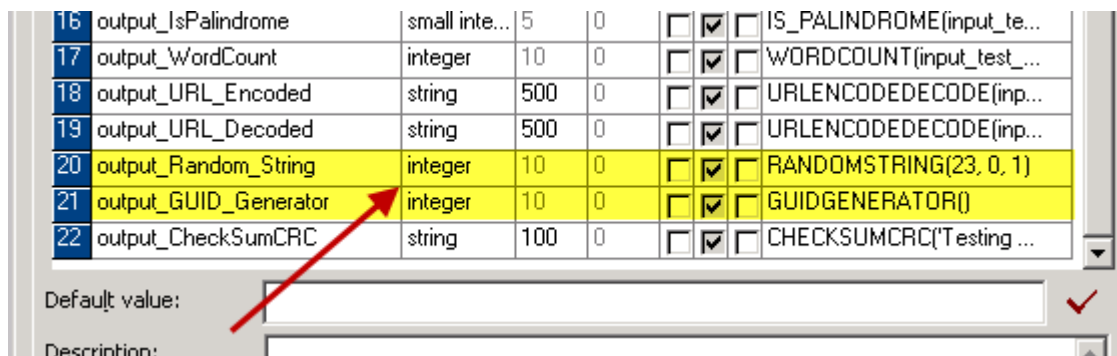
Several quickstart videos have been produced to educate developers on usage of this product. The quickstart videos are not a replacement for the core documentation found here and within the product directory. The quickstart videos can be found on the FyghtSoft website at <http://www.fyghtsoft.com/index.php/resources>.

INDIVIDUAL FUNCTION OPERATIONS

Informatica requires a separate binary file for each function. Before the plug-in is assigned in the Informatica Administrator Console the binary files should be present in the client and server bin folder locations on the respective physical server instances per license agreement.

RANDOM FUNCTIONS CAVEATS

Only two functions within the Premier Functions Toolbox version 1 fall within this category, RandomString and GUIDGenerator. These two functions need not accept any input ports as arguments and return a string value as their output. Although the RandomString function can access input ports values as an argument for dynamic use of the function, it does not have to. This non-dependence on an input causes Informatica to only recognize a numeric data type for the expression function. As seen in the screenshot below, ***please note that when using either function the data type should be set to integer***. Setting the data type to a string data type will result in a validation error.



URLENCODEDECODE

This function is an overload function that conducts the encoding or the decoding of string argument based on the second argument switch passed into the function. A typical use of this function would be to prepare a string for inclusion into a web URL to pass the universal web ASCII standard. The maximum input length of the first argument is 500 characters. The maximum output of any input string will be 1000 characters. The second argument should either be 'ENCODE' or 'DECODE' depending on the option sought.

IS_VALIDEMAILADDRESS

As a means to provide validate that an expected string or field holds the expected data integrity this function expects a string for certain characteristics following the global standard of RFC2821 and RFC2822. Basically this function will

approve all global email format structures based on this standard. ***Please note that this function does not verify an email address by pinging the mail server on which an email address may reside.***

CHECKSUMCRC

This function finally provides an explicit operation for conducting a well known data warehouse routine, checksum, within Informatica. As a caveat this function has the ability to accept an argument range as opposed to single argument as the out-of-the-box Informatica MD5 hash which is often used to conduct checksum operations. The range allows for between one (1) and fifty (50) arguments. Each argument can remain in its existing data type as no pre-conversion is required. In addition, if passing in input ports (columns) for dynamic use of this function all NULL values are handled within the function itself. Please note that NULL values are converted to "" (blank) within the core logic of the function, so if a table column contains both NULL and "" (blank) values the checksum has the chance of matching assuming all other columns/arguments passed into this function are the same.

CONCLUSION

The above functions are unique in that FyghtSoft is adding some new approaches to the way that transformation expressions have been utilized in the past. FyghtSoft aims to continue educating our customers about our products in such a way that transparency exists. If there are any questions, comments, or concerns please express these through our forums at <http://www.fyghtsoft.com/forums/> so that other users may also share in discovery and solution.