

# OpenText™ Actuate™ Big Data Analytics 5.2

OpenText Actuate Big Data Analytics 5.2 introduces several improvements that make the product more useful, powerful and flexible for end users. A new data loading module enables analysts and business users to load data directly into Big Data Analytics via a visual, drag-and-drop interface and offers built-in connections to more data sources. Powerful new expressions, based on regular expressions, now allow for complex finding and replacing of text in strings. And a new REST API enables developers to use Big Data Analytics results in applications across all devices.

## Table of Contents

Introduction.....	3
Data Loading Module .....	3
More Data Sources.....	4
Visual Data Loading .....	4
Data Engineering with Regular Expressions .....	4
REGEXMATCH.....	5
REGEXREPLACE .....	5
REST API.....	6
Upgrading to Big Data Analytics 5.2 .....	6
Supported Products .....	6

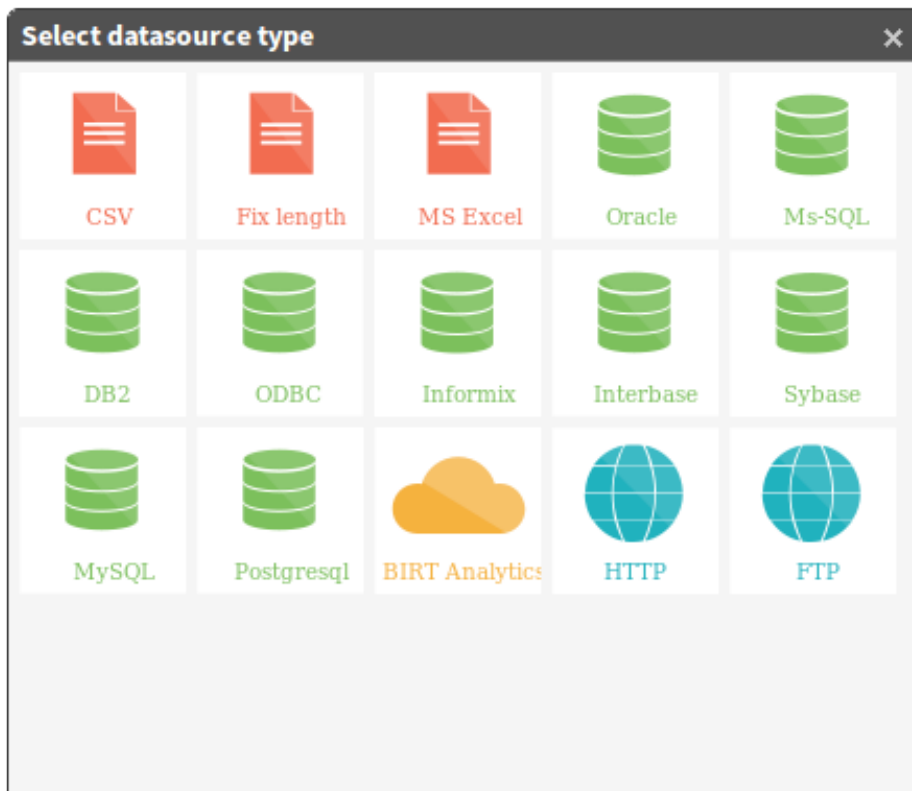
## Introduction

OpenText Actuate Big Data Analytics provides fast, free-form visual data mining and predictive analytics. Big Data Analytics combines easy-to-use data discovery and data mining tools with powerful and sophisticated analytic tools. To this foundation, Big Data Analytics 5.2 adds these improvements:

- A visual data loading module that handles more data types
- The addition of regular expressions for data enrichment in the engineering module
- A REST API that enables results from Big Data Analytics to be embedded in other applications

## Data Loading Module

A completely new module in Big Data Analytics provides an easier way to load data into its Fast DB. Found on a new tab between the home page and the explorer, the new module enables users to access data sources, decide what data to load from those sources, select which specific columns of data will be linked, and run the loading process—all using a graphical interface. Data loading is done within the Big Data Analytics application, and does not require using the qLoader ETL (extract, transform, load) module. The process to create a new connection is completely guided by the application and is accomplished with a few clicks. Users can add data sources without intervention or help from IT.



**FIGURE 1**

The initial screen of the new visual data loading module in OpenText Actuate Big Data Analytics

The process to create a new connection is completely guided by the application and is accomplished with a few clicks. Users can add data sources without intervention or help from IT.

### More Data Sources

The new module makes data sources accessible from the Big Data Analytics front end. Whereas previous releases only enabled users to load CSV and fixed-length flat files from the front end, the new greatly improved user interface supports more than a dozen data sources, including:

- Microsoft® Excel 97 through 2013
- Relational Databases, including
  - Oracle®
  - Microsoft® SQL Server®
  - IBM® DB2®
  - ODBC
  - MySQL®
  - PostgreSQL
- Other Big Data Analytics instances
- HTTP
- FTP

Big Data Analytics 5.2 introduces new options to load data from a Microsoft Excel spreadsheet or from another Big Data Analytics instance.

### Visual Data Loading

It has never been so easy to select, load, and link data sources for analysis in Big Data Analytics. Regardless of the data source, the load procedure is the same:

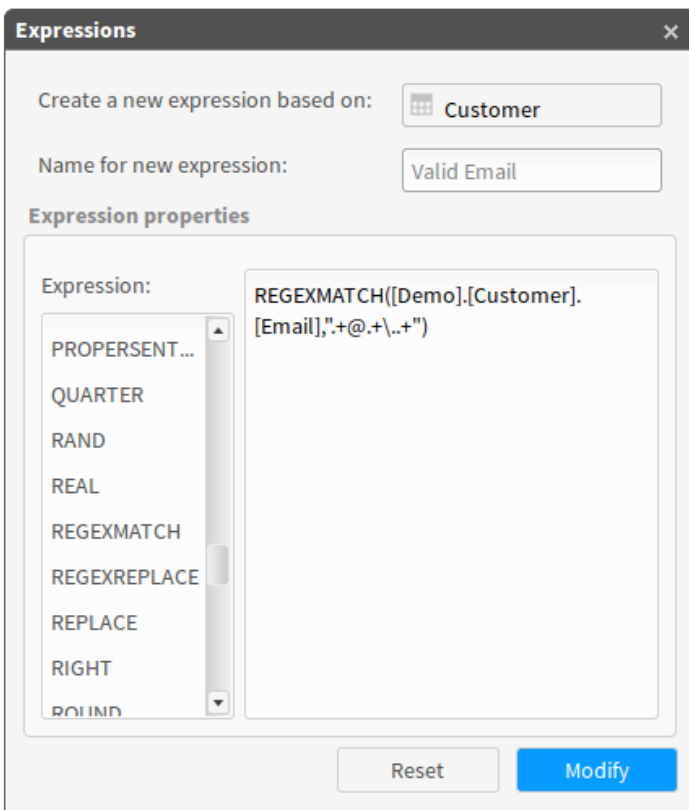
1. Create a connection to each source.
2. Define tables from those sources. Users also have the option of checking the data type, widths, and format.
3. Draw links between defined tables. Big Data Analytics will suggest the columns of each table to link, but the suggestions can be changed if the software links the wrong columns.
4. Load each table in Big Data Analytics' Fast DB. The software will create links between the tables.

## Data Engineering with Regular Expressions

Big Data Analytics includes new functions, known as *regular expressions*, for data enrichment, data quality assurance, and data filtering. Regular expressions are a commonly used technique for “find and replace” operations to check text strings, search for patterns in text, and rapidly make massive transformations. Big Data Analytics 5.2 comes with two new expressions that allow users to use regular expressions in engineering and enriching data:

- REGEXMATCH
- REGEXREPLACE

Both expressions can be called from the Engineering tab or from the qLoader module. Regular expressions are really powerful, but building a correct regular expression that solves a certain problem can be very complex. The user must have the ability to create the expression manually.



**FIGURE 2**

*Creating regular expressions in OpenText Actuate Big Data Analytics.*

### REGEXMATCH

The REGEXMATCH expression finds a certain pattern in a string value and responds with a Boolean result of 1 when the pattern exists in the string, or 0 when the pattern does not exist in the string.

One example of when REGEXMATCH is used is in classifying email addresses as well-formed or badly formed. This is often done as a precursor to analyzing if an email address is valid or not. The syntax for creating REGEXMATCH expressions is:

***REGEXMATCH("string field to find the pattern",regular\_expression)***

Big Data Analytics also offers the option to do searches that are not case-sensitive by adding "i" (for "insensitive").

***REGEXMATCH("string field to find the pattern",regular\_expression, "i")***

### REGEXREPLACE

The REGEXREPLACE expression offers the option of applying a replacement when the defined regular expression is found in the text string. The expression syntax is:

***REGEXREPLACE("string field to find the pattern", regular\_expression, replacement)***

As in REGEXMATCH, the REGEXREPLACE expression offers the option to do non-case-sensitive searches.

***REGEXREPLACE("string field to find the pattern", regular\_expression, replacement, "i")***

There are many examples of when regular expressions with the replace option are useful, but the most common situation is when a user wants to normalize certain data to fit a pattern – as is often done with telephone numbers, for example.

## REST API

The REST API in Big Data Analytics enables most results of analyses to be embedded in applications across all platforms, as well as in reports and dashboards created for OpenText™ Actuate™ Information Hub (iHub). The REST API is in addition to the existing Open Data Access (ODA) connector in Big Data Analytics; the purpose of the ODA is to display Big Data Analytics results in iHub reports and dashboards.

The REST API handles most basic analysis and advanced analysis within Big Data Analytics.

BASIC ANALYSIS HANDLED BY REST	ADVANCED ANALYSIS HANDLED BY REST
Crosstab	Forecast
Venn	Correlations
Profile	Linear Regression
Bubble	Logistic Regression
Pareto	Cluster
	Decision Tree

The REST API in Big Data Analytics employs the same look and feel as the REST API implementation in iHub, so developers who are familiar with REST in iHub will be able to use it immediately.

## Upgrading to Big Data Analytics 5.2

New installation of Big Data Analytics 5.2 requires a version-specific license key. More information about upgrading to Big Data Analytics 5.2 is available from OpenText Analytics sales representatives or by completing the form at <http://www.actuate.com/info/uscontactus/>

## Supported Products

Big Data Analytics includes support for the latest browsers, operating systems and databases. See the Supported Products Matrix for more information.

<http://developer.actuate.com/resources/supported-products/birt-analytics/5-2/>

[www.opentext.com](http://www.opentext.com)

NORTH AMERICA +800 499 6544 • UNITED STATES +1 847 267 9330 • GERMANY +49 89 4629-0  
UNITED KINGDOM +44 (0) 1189 848 000 • AUSTRALIA +61 2 9026 3400