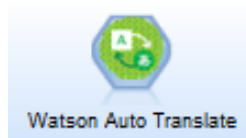


Watson Auto Translate

Use Watson Service through Bluemix to identify the language of a given text and translate it



Product: IBM® SPSS® Modeler

Extension type: Utility

Table of Contents

Description.....	3
Requirements.....	3
Installation.....	3
R Packages used.....	3
User Interface.....	4
Example.....	5
User Input.....	5-6
Results.....	7
Important links.....	7
Learn.....	7-8
Discuss.....	8

**Description:**

This node allows you to use the Watson Language Translation service on the Bluemix platform from within SPSS Modeler. This extension identifies the language of the input text and converts it to the desired target language. The output for this node is a data table that contains the original input data with additional new columns for the language of the input text and the translated text in the target language. Important: At this time, the Watson service cannot translate from non-English to non-English, meaning English must either be the source or target language.

Requirements:

- SPSS Modeler v16.0 or later
- R: <http://www.r-project.org/>
- 'R Essentials for SPSS Modeler' plugin: <https://developer.ibm.com/predictiveanalytics/downloads/>
- A Bluemix account with Watson Language Translation Service activated

Installation:

Close SPSS Modeler. Save the .cfe file in the CDB folder of the IBM SPSS Modeler installation directory for Windows and Linux. The copy should reside in that same folder and not in a sub-folder.

For example, for Windows 7 the default location is "C:\ProgramData\IBM\SPSS\Modeler\16\CDB". If the ProgramData folder is hidden type the path manually.

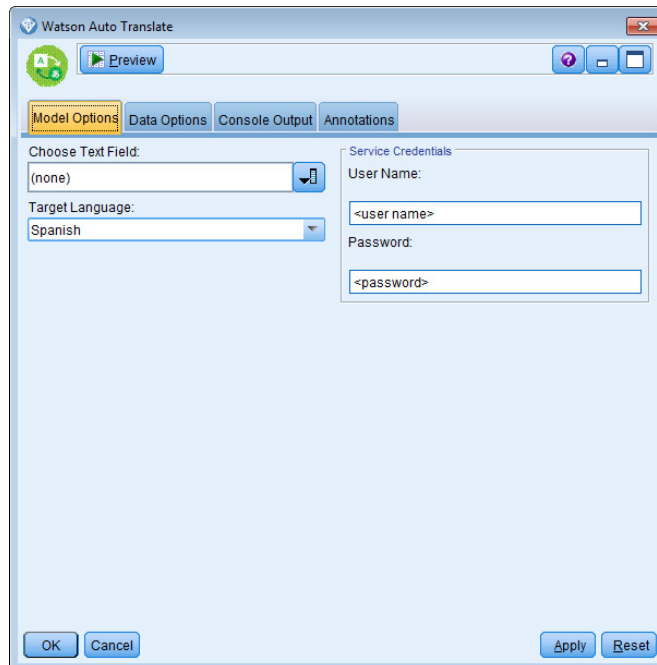
Restart SPSS Modeler: the node will now appear in the Field Ops palette.

R Packages used:

- 'httr' package created by Hadley Wickham Lang <https://cran.r-project.org/web/packages/httr/>
- 'XML' package created by Duncan Temple Lang <https://cran.r-project.org/web/packages/XML/>

User Interface

- Double click on the node to get to the options. There are the following fields:
 - Choose a Text Field: click the drop down to select the column containing the text that needs to be identified and translated
 - Target Language: This contains the list of languages that can be translated (English, Spanish, French, Portugese, Arabic)
 - User Name and Password: These are the credentials obtained in Bluemix when Watson Language Translation is added to your account (This is not your Bluemix user name and password).



Watson Auto Translate

Preview

Model Options | Data Options | Console Output | Annotations

Choose Text Field:
(none)

Target Language:
Spanish

Service Credentials

User Name:
<user name>

Password:
<password>

OK Cancel Apply Reset



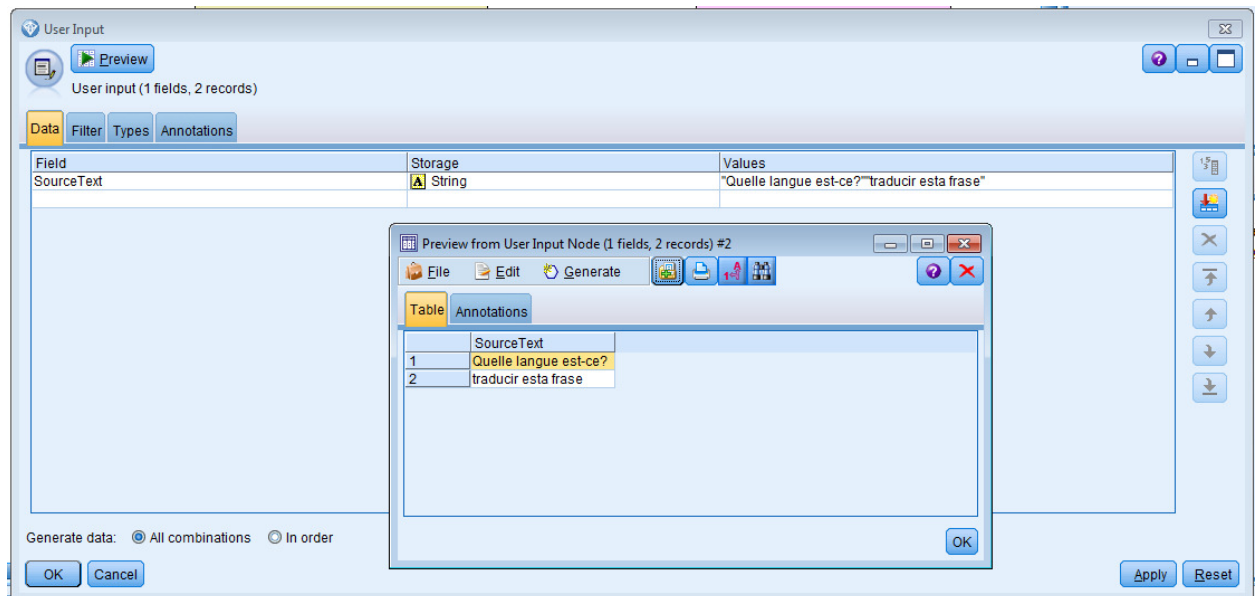
Example

This example will demonstrate how to identify and translate text from a User Input node. Text input can be used from any source node, the User Input node is used for this example for demonstration. For this example, the source text will be Spanish and French and the target language will be English.

User Input

1. In the sources palette, add a User Input node to the stream
2. Name the first Field “SourceText” and enter the following text in the Values column:

"Quelle langue est-ce?" "Traducir esta frase"

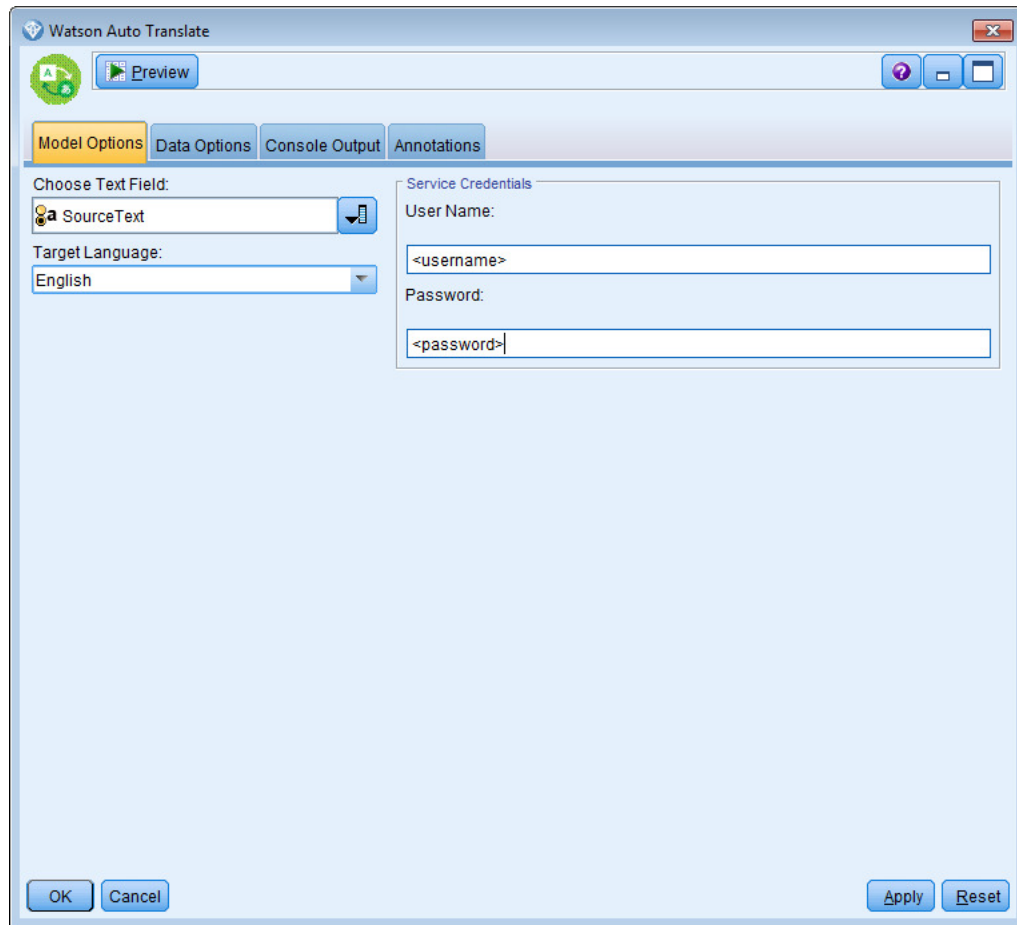


3. Add the Watson Auto Translate node to the stream and connect it to the User Input
4. Now fill the fields:

Choose Text Field: Click the drop down and pick “Source Text”

Target Language: Choose English for this example

User Name and Password: Enter the credentials obtained by adding Watson Translation Service to your Bluemix Dashboard



5. From the Output palette, connect a Table to the translation node
6. Click on the Table and run the stream





Results

If you used the sample text above, your results should match the table below:

	SourceText	Source Language ID	Translated Text
1	Quelle langue est-ce?	fr	What language is that?
2	traducir esta frase	es	Translate this phrase

Notice that two columns were added to the original User Input data. The Source Language ID is the two letter identifier for the source text. The translated text is the translation from the source language to the target language of English. The table below can be used as a reference for the Source Language ID column:

Language	Code
English	En
Spanish	Es
French	Fr
Portuguese	Pt
Arabic	Ar

Important Links

Learn

- Learn more about [SPSS software](#).
- Visit [developerWorks Business analytics](#) for more technical analytics resources for developers.
- Create your Bluemix account here: <https://console.ng.bluemix.net/solutions/watson>
- The [Comprehensive R Archive Network](#) is the main site for the R project and each R package. The help pages and manuals that are associated with optimx, nlrmr, and Rcgmin are detailed. Numerous references are provided.



- Read "[Do I need to learn R?](#)" (Catherine Dalzell, developerWorks, September 2013) to learn why R is a valuable tool for data analytics that was expressly designed to reflect the way that statisticians think and work.
- "[Calling R from SPSS](#)" describes how to use R code inside IBM SPSS Modeler 16.
- Read "[Create new nodes for IBM SPSS Modeler 16 using R](#)" to learn how to create new extensions easily.

Discuss

- Visit the [IBM SPSS Community](#) to share tips and experiences with other IBM SPSS developers.
- Follow [developerWorks on Twitter](#) to be among the first to hear about new resources.