NLU

A Watson Explorer Content Analytics UIMA analysis engine for Watson Natural Language Understanding.

Revision History

Date	Version	Status	Description	Author
06/04/2017	1.0	Release	Initial	Martin Saunders

Contents

1. Purpose ..... Error! Bookmark not defined.
2. Installation 4
2.1 Pre-Requisites 4
2.2 Content Analytics Studio 4

1. Introduction

The Watson Natural Language Understanding service uses natural language processing to analyze semantic features of any text. This NLU annotator allows those identified semantic features to be annotated and utilised in any UIMA compliant pipeline such as Watson Explorer Content Analytics server to Content Analytics Studio.

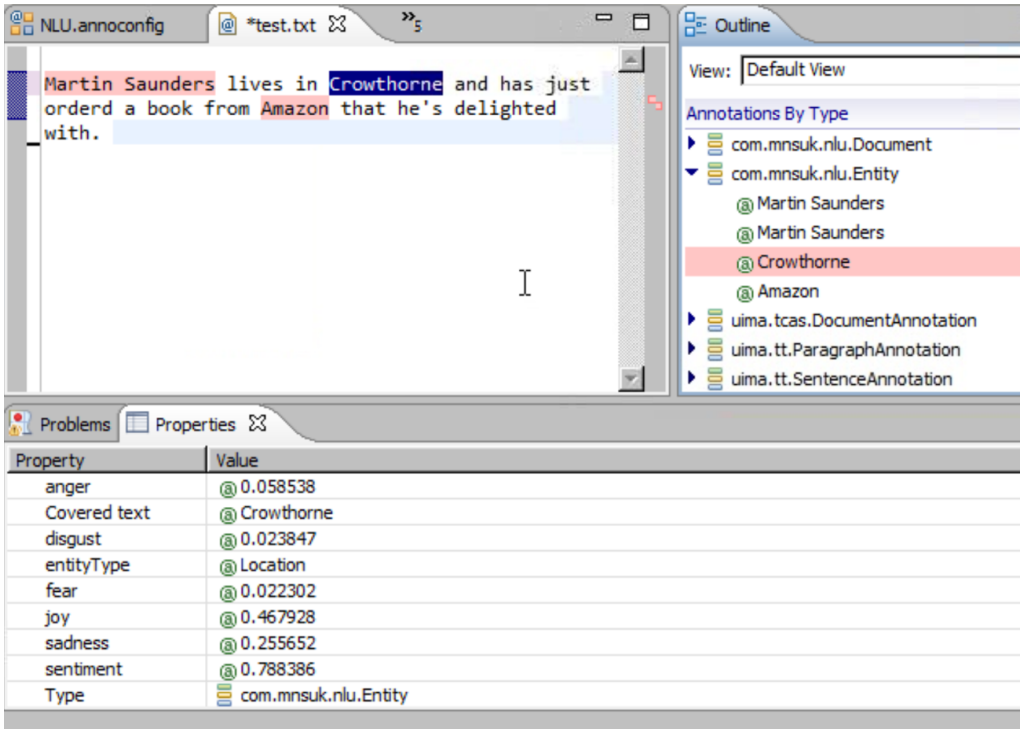
It supports the the following features of the Watson NLU service

- Entities
  Identify people, cities, organisations and many other entity types in the text
  - Sentiment analysis on identified entities
  - Emotion analysis on identified entities

- Sentiment
  Analyse the general sentiment of the document

- Emotion
  Detect emotion conveyed by the entire document.

For example the screen shot below shows an Entity annotation with targeted sentiment and emotion enabled.



1.1 Languages Supported

The annotator supports the same languages as the Watson NLU service. At the time of writing these are:

Entities: Arabic, English, French, German, Italian, Portuguese, Russian, Spanish and Swedish.
Emotion: English only
Sentiment: Arabic, English, French, German, Italian, Portuguese, Russian and Spanish.

2. Installation

2.1 Pre-Requisites

- Installation of Content Analytics Studio version 11 or above
- NLU files:
  - Core files (included in distribution):
    - NLU-ae.xml
    - NLU-n-n-n.jar
    - ICAUIMAUtils-n-n-n.jar
  - Watson Java JDK
    - core-n.n.n.jar
    - natural-language-understanding-n.n.n.jar
    - okhtpp-n.n.n.jar
    - okhttp-urlconnection-n.n.n.jar
    - okio-n.n.n.jar
    - gson-n.n.n.jar
    - logging-interceptor-n.n.n.jar
- Credentials for an instance of the Watson NLU service deployed in Bluemix.

2.2 Content Analytics Studio

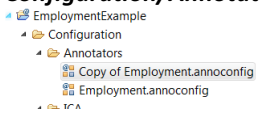
2.2.1 Create a project.

2.2.2 In this project create a folder to hold the NLU resources. This can be anywhere in the workspace but something like ***Resources/Custom/NLU*** would be a good choice.

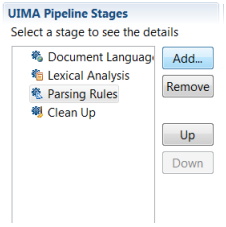
2.2.3 Copy the annotation engine configuration file (***NLU-ae.xml***) and all the jar files that make up NLU into this new folder.

2.2.4 In the relevant UIMA pipeline configuration file in your project add a custom stage as the penultimate stage of the pipeline. To do this:

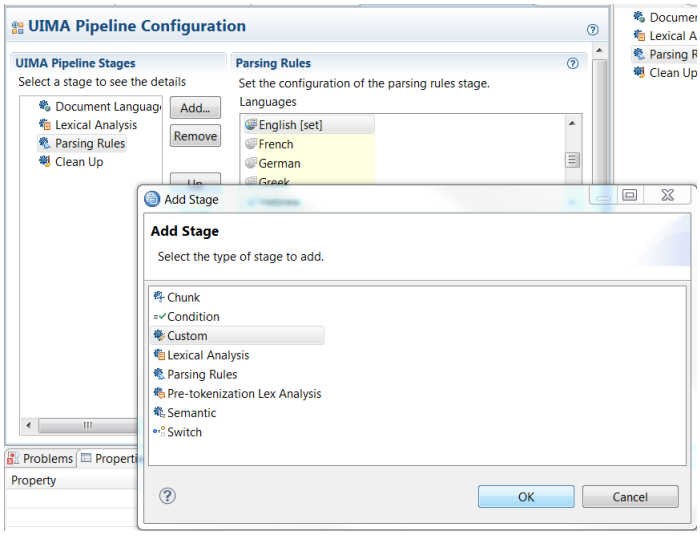
- Double click on the relevant pipeline configuration file under your ***Configuration/Annotators*** folder to open and edit it.



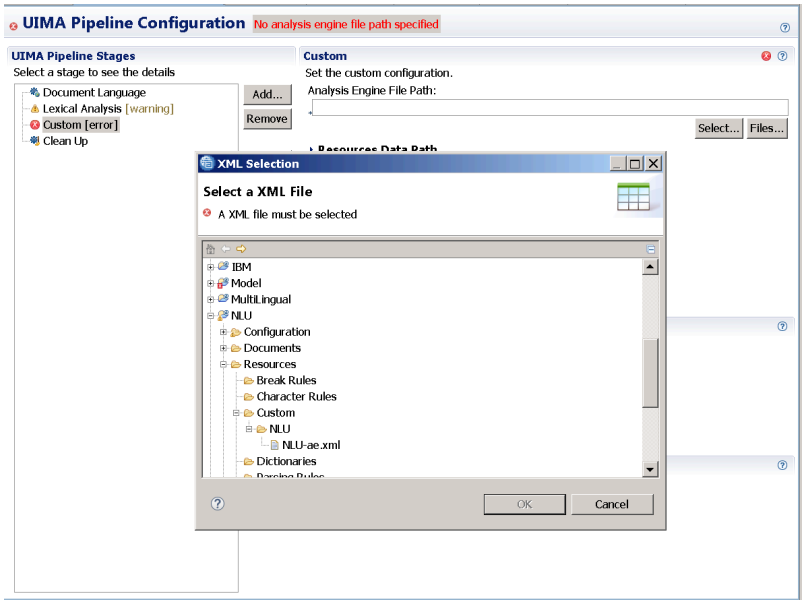
- select the Lexical Analysis stage
  click the “Add” button



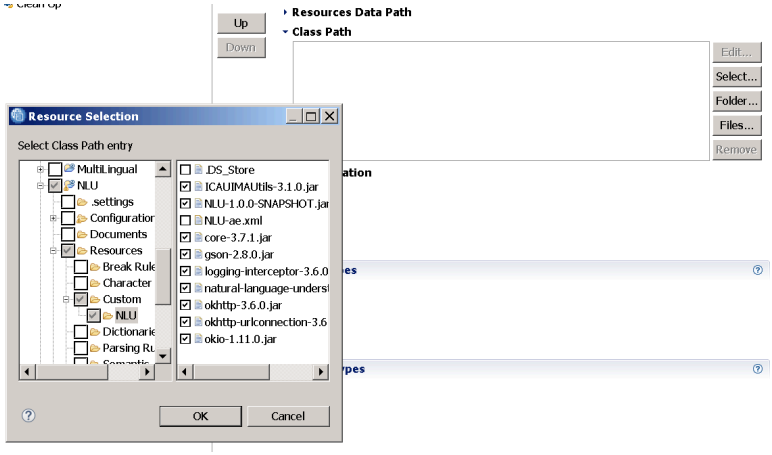
- select “Custom” in the popup window
  click “OK”



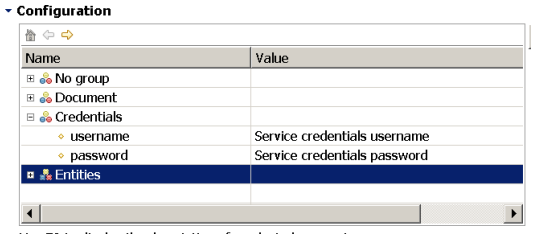
- Click in the text box of the Custom panel under where it says ***Analysis Engine File Path:*** and click the ***Select*** button. Navigate to the folder created in step 2.2.2 and select the ***NLU-ae.xml*** annotation engine configuration file.



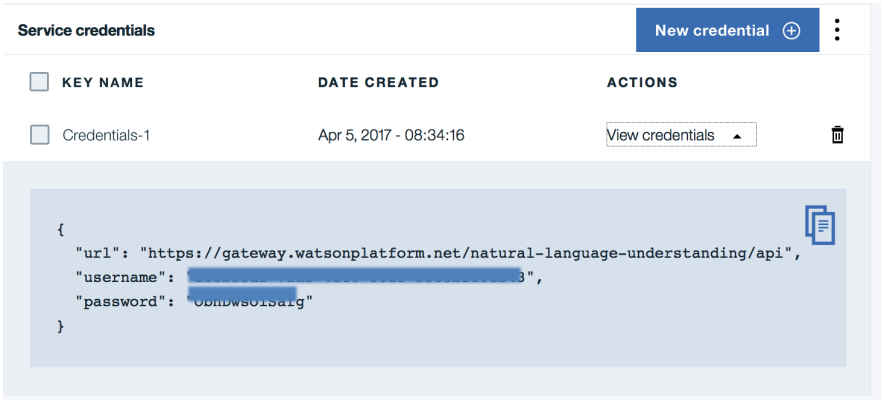
- Back in the ***Custom*** panel click the small right arrow to open the ***Class Path*** panel and click ***Select***. Again navigate to the new folder and this time click the check boxes to select all the jar files.
  Click ***OK***



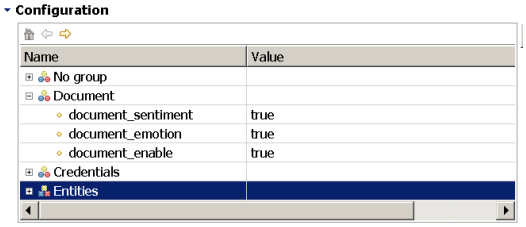
- Back in the ***Custom*** panel click the small right arrow to open the ***Configuration*** panel, then click the small right arrow to open the ***Credentials*** Configuration group.



Set the username and password with the credentials from your Watson NLU service deployed in Bluemix.



- Close the ***Credentials*** group and then click to open the ***Document*** group,

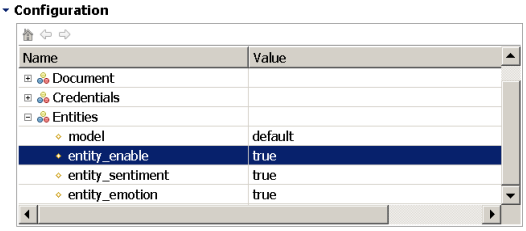


Set the parameters according to the guidance in the table below.

Parameter	Default value	Note
document_enable	true	true | false Enable document level analysis.
document_sentiment	true	true | false Set this to true to enable document level sentiment analysis.
document_emotion	true	true | false Set this to true to enable document level emotion analysis.

Note: if ***document_enable*** is set to "true" then at least one of the other parameters must also be set to true.

- Close the ***Document*** group and click the right arrow to open the ***Entities*** group,



Set the parameters according to the guidance in the table below.

Parameter	Default value	Note
entity_enable	true	true | false Enable entity level analysis on the document
entity_sentiment	true	true | false Set this to true to enable targetted sentiment analysis for detected entities.
entity_emotion	true	true | false Set this to true to enable targetted emotion analysis for detected entities.
model	default	NOT IMPLEMENTED YET Valid model name for a deployed custom WKS model.

- Save the UIMA pipeline configuration file
- You're ready to go!