# Rosette ES Plugin Test

### Test environment

|  |  |
| --- | --- |
| os | Linux ip-10-152-12-219 4.14.62-65.117.amzn1.x86\_64 #1 SMP Fri Aug 10 20:03:52 UTC 2018 x86\_64 x86\_64 x86\_64 GNU/Linux |
| elasticsearch | Elasticsearch-6.1.2 |
| Rosette Plugin | rbl-je-elasticsearch-6.1.2.1.zip |
| Node private ip | 10.152.12.219 |

### Installation

Navigate to the elasticsearch-6.1.2 root directory and run:

ssh -i KP\_infrastructure.pem ec2-user@10.152.12.219 -oUserKnownHostsFile=/dev/null -oStrictHostKeyChecking=no

cd /opt/reuters/data/elasticsearch/unit\_test/elasticsearch-6.1.2

bin/elasticsearch-plugin install file:/path/to/rbl-je-elasticsearch-6.1.2.1.zip

cp ./rlp-license.xml ./plugins/analysis-rbl-je/rbl-je-7.23.2.c59.0/licenses/

./bin/elasticsearch -d -p PID

### Arabic analyzer settings

"rbl\_ara": {

"type": "rbl",

"language" : "ara",

"stopwords" : "\_arabic\_"

},

### Create an index for testing Arabic analzyer

curl -XPUT -H "Content-Type: application/json" http://localhost:9230/arci-test --data @arci.mapping.json

### Test case

Case 1: “**تنقلين**”, a verb in Arabic, rbl plugin can return this word’s dictionary form normally

curl -XPOST -H "Content-Type: application/json" http://localhost:9230/arci-test/\_analyze?pretty -d '{

"analyzer":"rbl\_ara",

"text": "تنقلين"

}'

{

"tokens" : [

{

"token" : "تنقلين",

"start\_offset" : 0,

"end\_offset" : 6,

"type" : "<ALNUM>",

"position" : 0

},

**{**

**"token" : "نَقَل",**

**"start\_offset" : 0,**

**"end\_offset" : 6,**

**"type" : "<LEMMA>",**

**"position" : 0**

**}**

]

}

Case 2 “**أمنحي**”, a verb in Arabic, rbl plugin doesn’t return its dictionary form.

curl -XPOST -H "Content-Type: application/json" http://localhost:9230/arci-test/\_analyze?pretty -d '{

"analyzer":"rbl\_ara",

"text": "أمنحي"

}'

{

"tokens" : [

{

"token" : "أمنحي",

"start\_offset" : 0,

"end\_offset" : 5,

"type" : "<ALNUM>",

"position" : 0

}

]

}

Case 3 **“كنت”** a verb in Arabic, rbl plugin doesn’t return its dictionary form.

curl -XPOST -H "Content-Type: application/json" http://localhost:9230/arci-test/\_analyze?pretty -d '{

"analyzer":"rbl\_ara",

"text": "كنت"

}'

{

"tokens" : [

{

"token" : "كنت",

"start\_offset" : 0,

"end\_offset" : 3,

"type" : "<ALNUM>",

"position" : 0

}

]

}

Case 4 **“سياراتهنا”** a noun in Arabic, rbl plugin doesn’t return its dictionary form.

curl -XPOST -H "Content-Type: application/json" http://localhost:9230/arci-test/\_analyze?pretty -d '{

"analyzer":"rbl\_ara",

"text": "سياراتهنا"

}'

{

"tokens" : [

{

"token" : "سياراتهنا",

"start\_offset" : 0,

"end\_offset" : 9,

"type" : "<ALNUM>",

"position" : 0

}

]

}

Case 5 **“علمكنا”,** this word is derived from **“علم”,** rbl plugin doesn’t return its dictionary form.

curl -XPOST -H "Content-Type: application/json" http://localhost:9230/arci-test/\_analyze?pretty -d '{

"analyzer":"rbl\_ara",

"text": "علمكنا"

}'

{

"tokens" : [

{

"token" : "علمكنا",

"start\_offset" : 0,

"end\_offset" : 6,

"type" : "<ALNUM>",

"position" : 0

}

]

}

Case 6 **“علما”,** a verb in Arabic, rbl plugin can return its dictionary form normally

curl -XPOST -H "Content-Type: application/json" http://localhost:9230/arci-test/\_analyze?pretty -d '{

"analyzer":"rbl\_ara",

"text": "علما"

}'

{

"tokens" : [

{

"token" : "علما",

"start\_offset" : 0,

"end\_offset" : 4,

"type" : "<ALNUM>",

"position" : 0

},

**{**

**"token" : "عِلْم",**

**"start\_offset" : 0,**

**"end\_offset" : 4,**

**"type" : "<LEMMA>",**

**"position" : 0**

**}**

]

}

But when I type this word in the [TRY THE DEMO](https://www.rosette.com/capability/base-linguistics/#try-the-demo) in Rosette website, it returns none.

