

## MLSS Tagger Web service

### 1 BASIC INFORMATION

#### 1.1 Overview and purpose of the tool

The part of speech tagger for Maltese is based on TnT, the statistical part of speech tagger by Thorsten Brants (<http://www.coli.uni-saarland.de/~thorsten/tnt/>). It was modified for the Maltese Language Resource Server (MLRS) by Albert Gatt (Linguistics Department, University of Malta). The model for Maltese was trained on manually tagged texts and has reached an accuracy of 96%. Below is a list of the tags which are used, along with a description. The web service for the Maltese Language Software Services (MLSS) was provided by Andrew Attard.

#### 1.2 Tag list

<b>CC</b>	coordinating conjunction	<b>PI</b>	pronoun, indefinite
<b>CMP</b>	complementiser	<b>PMP</b>	preposition ma' with bound pronoun
<b>CS</b>	subordinating conjunction	<b>PP</b>	pronoun, personal
<b>DD</b>	determiner	<b>PR</b>	pronoun, reflexive
<b>DDC</b>	definite determiner, clitic	<b>PRP</b>	preposition
<b>DP</b>	determiner, plural	<b>PRPC</b>	fused preposition-article
<b>DQ</b>	determiner quantifier	<b>PT</b>	pronoun, possessive
<b>DS</b>	specifier, singular	<b>PUN</b>	punctuation
<b>EX</b>	existential marker	<b>RA</b>	residual, acronym
<b>II</b>	interjection	<b>RB</b>	residual, abbreviation
<b>MJ</b>	modifier, adjective	<b>RD</b>	residual, date
<b>MV</b>	modifier, adverb	<b>RFR</b>	residual, formula, mathematical symbol
<b>NCI</b>	numeral, cardinal, intransitive	<b>RFW</b>	residual, foreign word
<b>NCT</b>	numeral, cardinal, transitive	<b>RH</b>	residual, honorific
<b>NN</b>	common noun	<b>RO</b>	residual, other
<b>NO</b>	numeral, ordinal	<b>RS</b>	residual, other symbol
<b>NP</b>	proper name	<b>UAM</b>	(unique,unassigned) multiword utterance
<b>NPI</b>	initial in proper name	<b>VA</b>	verb, auxiliary
<b>NV</b>	verbal negator	<b>VG</b>	pseudo verb
<b>PAC</b>	particle, aspect marker, continuous aspect	<b>VP</b>	participle, active, or passive
<b>PAF</b>	particle, aspect marker,	<b>VV</b>	main verb

	prospective aspect	
PD	pronoun, demonstrative	

### 1.3 Using the tagger

The WSDL link is <http://metanet4u.research.um.edu.mt/services/MtPOS?wsdl>.

The service has two methods which can be invoked:

- String tagOneWordReturn(String text)
- String tagParagraphReturn(String text)

Both methods take a string as input, i.e., the text to be tagged, and return another string, i.e., the tagged text. The difference between the two is tagOneWordReturn returns the output as one word per line, while the other returns it as tagged paragraphs (if any paragraphs occur in the input string).

The format of the output is:

word\_TAG

## 2 TECHNICAL INFORMATION

### 2.1 WSDL description

This is available at <http://metanet4u.research.um.edu.mt/services/MtPOS?wsdl> and is given below:

```
<?xml version="1.0" encoding="UTF-8"?>
<wsdl:definitions targetNamespace="http://postagger.mt.edu.um"
xmlns:apachesoap="http://xml.apache.org/xml-soap"
xmlns:impl="http://postagger.mt.edu.um"
xmlns:intf="http://postagger.mt.edu.um"
xmlns:wsdl="http://schemas.xmlsoap.org/wsdl/"
xmlns:wsdlsoap="http://schemas.xmlsoap.org/wsdl/soap/"
xmlns:xsd="http://www.w3.org/2001/XMLSchema">
<!--WSDL created by Apache Axis version: 1.4
Built on Apr 22, 2006 (06:55:48 PDT)-->
<wsdl:types>
  <schema elementFormDefault="qualified"
targetNamespace="http://postagger.mt.edu.um"
xmlns="http://www.w3.org/2001/XMLSchema">
    <element name="tagOneWordReturn">
      <complexType>
        <sequence>
          <element name="text" type="xsd:string"/>
        </sequence>
      </complexType>
    </element>
    <element name="tagOneWordReturnResponse">
      <complexType>
        <sequence>
          <element name="tagOneWordReturnReturn"
type="xsd:string"/>
        </sequence>
      </complexType>
    </element>
    <element name="tagParagraphReturn">
      <complexType>
        <sequence>
          <element name="text" type="xsd:string"/>
        </sequence>
      </complexType>
    </element>
  </schema>
</wsdl:types>
```

```

        </sequence>
    </complexType>
</element>
<element name="tagParagraphReturnResponse">
    <complexType>
        <sequence>
            <element name="tagParagraphReturnReturn"
type="xsd:string"/>
        </sequence>
    </complexType>
</element>
</schema>
</wsdl:types>

    <wsdl:message name="tagOneWordReturnRequest">

        <wsdl:part element="impl:tagOneWordReturn"
name="parameters">

            </wsdl:part>

        </wsdl:message>

    <wsdl:message name="tagParagraphReturnResponse">

        <wsdl:part element="impl:tagParagraphReturnResponse"
name="parameters">

            </wsdl:part>

        </wsdl:message>

    <wsdl:message name="tagParagraphReturnRequest">

        <wsdl:part element="impl:tagParagraphReturn"
name="parameters">

            </wsdl:part>

        </wsdl:message>

    <wsdl:message name="tagOneWordReturnResponse">

        <wsdl:part element="impl:tagOneWordReturnResponse"
name="parameters">

            </wsdl:part>

        </wsdl:message>

    <wsdl:portType name="MtPOS">

        <wsdl:operation name="tagOneWordReturn">

            <wsdl:input message="impl:tagOneWordReturnRequest"
name="tagOneWordReturnRequest">

```

```

        </wsdl:input>

        <wsdl:output message="impl:tagOneWordReturnResponse"
name="tagOneWordReturnResponse">

        </wsdl:output>

</wsdl:operation>

<wsdl:operation name="tagParagraphReturn">

        <wsdl:input message="impl:tagParagraphReturnRequest"
name="tagParagraphReturnRequest">

        </wsdl:input>

        <wsdl:output message="impl:tagParagraphReturnResponse"
name="tagParagraphReturnResponse">

        </wsdl:output>

</wsdl:operation>

</wsdl:portType>

<wsdl:binding name="MtPOSSoapBinding" type="impl:MtPOS">

        <wsdlsoap:binding style="document"
transport="http://schemas.xmlsoap.org/soap/http"/>

        <wsdl:operation name="tagOneWordReturn">

                <wsdlsoap:operation soapAction=""/>

                <wsdl:input name="tagOneWordReturnRequest">

                        <wsdlsoap:body use="literal"/>

                </wsdl:input>

                <wsdl:output name="tagOneWordReturnResponse">

                        <wsdlsoap:body use="literal"/>

                </wsdl:output>

        </wsdl:operation>

        <wsdl:operation name="tagParagraphReturn">

                <wsdlsoap:operation soapAction=""/>

                <wsdl:input name="tagParagraphReturnRequest">

                        <wsdlsoap:body use="literal"/>

                </wsdl:input>

```

```

        <wsdl:output name="tagParagraphReturnResponse">
            <wsdlsoap:body use="literal"/>
        </wsdl:output>
    </wsdl:operation>
</wsdl:binding>
<wsdl:service name="MtPOSService">
    <wsdl:port binding="impl:MtPOSSoapBinding" name="MtPOS">
        <wsdlsoap:address
location="http://metanet4u.research.um.edu.mt/services/MtPOS"/>
    </wsdl:port>
</wsdl:service>
</wsdl:definitions>

```

## 2.2 Software dependencies and system requirements

Not applicable.

## 2.3 Execution instructions

For how to invoke the web service from your own application, see the WSDL above.

## 2.4 Input/Output data formats

Input data format: text string

Output data format: a text string with the tagged words in the format **word\_TAG**, either as one word per line, or as tagged paragraphs.

## 2.5 Integration with external tools

The web service does not need any external tool.

## 3 CONTENT INFORMATION

An input string such as

Il-kunċett ta' Malta bħala destinazzjoni għall-Cruise Liners għandha tingħata spinta akbar.

will result in the following output strings.

**Paragraph return:**

Il-\_DDC kunċett\_NN ta'\_PRP Malta\_NP bħala\_MV destinazzjoni\_NN għall-  
 \_PRPC Cruise\_NP Liners\_NP għandha\_VG tingħata\_VV spinta\_NN akbar\_MJ  
 .\_PUN

**One word return:**

Il-\_DDC  
 kunċett\_NN  
 ta'\_PRP  
 Malta\_NP  
 bħala\_MV  
 destinazzjoni\_NN  
 għall-\_PRPC  
 Cruise\_NP  
 Liners\_NP  
 għandha\_VG

tinghata\_VV  
spinta\_NN  
akbar\_MJ  
.\_PUN

#### 4 ADMINISTRATIVE INFORMATION

##### *4.1 Contact*

For further information and technical support running the webservice, please send an email to Andrew

Attard: [andrew.attard@um.edu.mt](mailto:andrew.attard@um.edu.mt)

For information on the MLRS tagger, please send an email to Albert Gatt: [albert.gatty@um.edu.mt](mailto:albert.gatty@um.edu.mt)

##### *4.2 Delivery medium (if relevant; description of the content of each piece of medium)*

Web service

##### *4.3 Copyright statement and information on IPR*

META-SHARE Commons BY-NC-SA

#### 5 RELEVANT REFERENCES AND OTHER INFORMATION

<http://metanet4u.research.um.edu.mt/>

<http://www.coli.uni-saarland.de/~thorsten/tnt/>

<http://mlrs.research.um.edu.mt/>