

# **Technical Description of the Named Entities Recognition web service**

## **Version 1.2**

---

## Content

1. Modification .....	3
Version 1.1 .....	3
Version 1.2 .....	3
2. Overview .....	3
3. Web-service access.....	3
3.1. SOAP access .....	3
3.2. REST access.....	4
4. Function description.....	4
Synapse_Analysis_Output_NamedEntities GetNamedEntities (Synapse_Analysis_Input input); .....	4
5. Consumption example.....	6

---

## 1. Modification

### **Version 1.1**

- Unified credentials system.

### **Version 1.2**

- Microdata output format

## 2. Overview

The aim of this document is to present and explain the Named Entities Recognition Web-service API. It provides a description of the web-service function, and a comprehensive example of consumption.

## 3. Web-service access

This web-service is available both through a SOAP and a REST access point. This section describes how to connect to one of these two access points.

If you have been provided a specific URL for your project, please replace [client] in the URL with your client identifier. If not, just remove the "[client]" from the URL.

### **3.1. SOAP access**

The web-service expose a WSDL at this address:

<http://api-synapse-dev.azurewebsites.net/namedentities/NamedEntitiesWCF.svc?wsdl>

This WSDL can be used to automatically generate a proxy for your application.

#### **Manual configuration**

If you cannot use the WSDL, or if you prefer to build manually the request, here are the connection parameters for the service:

Service address:

<http://api-synapse-dev.azurewebsites.net/namedentities/NamedEntitiesWCF.svc>

Service host:

<http://api-synapse-dev.azurewebsites.net/namedentities>

SOAPAction:

---

<http://syn-web01.synapse-fr.com/api/namedentities/INamedEntitiesWCF/GetNamedEntities>

The web-service is configured to receive and send XML.

### 3.2. REST access

In order to provide a convenient way to call the web service from a web interface, the function is also available at:

<http://api-synapse-dev.azurewebsites.net/namedentities/NamedEntitiesWCF.svc/rest/GetNamedEntities>

#### **Important**

When using the REST access point, one should check that the parameters are given in the alphabetic order (order given below, with *password* before *user* name). The web-service may ignore some of the parameters if they are not called in the proper order.

The section 6 provides an example of how to call the web-service using JQuery and the REST access point.

## 4. Function description

### **Synapse\_Analysis\_Output\_NamedEntities GetNamedEntities (Synapse\_Analysis\_Input input);**

#### **Description:**

The function takes as input a structure containing the credentials identifying the client willing to consume the web-service, the language of the text, the text to analyse and the desired output format.

The function returns another structure with the identified named entities in the specified format and the possible exception and error codes raised by the service.

#### **Argument:**

*input* is a structure containing the following elements:

- *credentials*: a structure containing
  - o *password*: the corresponding password (as given by Synapse)
  - o *user*: the username (as given by Synapse)
- *format*: (optional) desired output format. Default is “weblab”. Two formats are supported: “weblab” and “microdata”.
- *lang*: language of the text. Can take the values “fr” for French, and “en” for English.
- *text*: the text to be processed

XML example for input:

```
<Synapse_Analysis_Input>
  <credentials>
    <password>password</password>
    <user>client</user>
```

```
</credentials>
<format>microdata</format>
<lang>fr</lang>
<text> La Fédération a annoncé jeudi dernier dans un communiqué que
l'éthiopien Kenenisa Bekele et le jamaïcain Usain Bolt ont été désignés
parmi les cinq finalistes pour le titre d'athlète de l'année
2013..</text>
</Synapse_Analysis_Input>
```

**Returns:**

*output* is a structure containing the following elements:

- *namedEntities*: result of the named entity identification serialized in the specified format. In its current version, the service identifies persons, organizations, places and dates.
- *exceptions*: structure containing
  - o *errcode*: the error code. 0 if no error were raised.
  - o *message*: comprehensive message describing the problem if an error was raised (i.e. if the errcode is different from 0).

Typical XML output corresponding to the input example (some extra concept and keyword nodes pruned for lisibility):

```
<?xml version="1.0" encoding="utf-16"?>
<Synapse_Analysis_Output_NamedEntities
xmlns="http://schemas.datacontract.org/2004/07/NamedEntities_WCF"
xmlns:i="http://www.w3.org/2001/XMLSchema-instance">
  <exceptions
xmlns="http://schemas.datacontract.org/2004/07/Synapse.CommonWS">
    <errcode>0</errcode>
    <message i:nil="true" />
  </exceptions>
  <namedEntities>La &lt;span itemscope
itemtype='http://schema.org/Organization'&gt;&lt;span
itemprop='name'&gt;Fédération&lt;/span&gt;&lt;/span&gt; a annoncé
&lt;date&gt;jeudi dernier&lt;/date&gt; dans un communiqué que
l'éthiopien &lt;span itemscope
itemtype='http://schema.org/Person'&gt;&lt;span
itemprop='name'&gt;Kenenisa Bekele&lt;/span&gt;&lt;/span&gt; et le
jamaïcain &lt;span itemscope
itemtype='http://schema.org/Person'&gt;&lt;span
itemprop='name'&gt;Usain Bolt&lt;/span&gt;&lt;/span&gt; ont été
désignés parmi les cinq finalistes pour le titre d'athlète &lt;date
datetime='2013'&gt;de l'année 2013&lt;/date&gt;.
  </namedEntities>
</Synapse_Analysis_Output_NamedEntities>
```

Error code values (errcode):

- 0: no error.
- 1: wrong credentials.
- 2: expired credentials, user does no longer have the right to consume web-service. Contact Synapse to renew your subscription.

- 3: unrecognized language – analysis returned for the default language.
- 4: no text specified, or unrecognized text format.
- 5: internal error. If problem persist, contact support.
- 6 unrecognized output format – analysis returned in default format.

## 5. Consumption example

This section shows a minimal html page exploiting the web service.

```
<!DOCTYPE html>
<html>
<head>
  <title>Test Named Entities Identification Web Service</title>
  <style>
    span, date {
      font-weight: bold;
    }
  </style>
</head>

<body>
  <h1>Test Named Entities Identificatino Web Service</h1>

  <div id="result"></div>

  <script
src="http://ajax.googleapis.com/ajax/libs/jquery/1.11.2/jquery.min.js"></sc
ript>
  <script>window.jQuery || document.write('<script
src="js/vendor/jquery-1.11.2.min.js"></script>')</script>

  <script>
var pass = "your password";
var user = "your username";
var text = "La Fédération a annoncé jeudi dernier dans un communiqué que
l'éthiopien Kenenisa Bekele et le jamaïcain Usain Bolt ont été désignés
parmi les cinq finalistes pour le titre d'athlète de l'année 2013.";
var xmlInput = '<Synapse_Analysis_Input><credentials><password>' + pass +
'</password><user>' + user +
'</user></credentials><date></date><format>microdata</format><lang>fr<
/lang><text>' + text + '</text></Synapse_Analysis_Input>';

function parseResponse(output_analysis) {
  var annotatedText = $(output_analysis).find('namedEntities').text();
  $("#result").html(annotatedText);
}

function errorHandler(XHR, textStatus, errorThrown) {
  alert(errorThrown);
}

$.ajax({
  type : "POST",
```

```
url : "http://api-synapse-  
dev.azurewebsites.net/namedentities/NamedEntitiesWCF.svc/rest/GetNamedEntit  
ies",  
  data : xmlInput,  
  dataType : 'xml',  
  contentType : "text/xml",  
  success : parseResponse,  
  error : errorHandler  
});  
  </script>  
</body>  
</html>
```