

NERITS

# A Machine Translation mashup system using Wikimeta and LOD

Kamel Nebhi, Luka Nerima, Eric Wehrli

Laboratory of Analysis and Technology of Language  
University of Geneva, Switzerland

AI Mashup Challenge 2013



# Motivation

- Machine Translation (MT) has become increasingly commonplace in everyday use through BabelFish, Bing Translator or Google Translate.
- MT approaches provide good results but none of them exploit contextual information like named entity to help user comprehension.

## Solution

- Named Entity Recognition for the Interactive Translation System
- **The goal** is to propose a cross-lingual translation by providing detailed information extracted from DBpedia about persons, locations and organizations in the mother tongue of the user.
- This helps at scaling the traditional multilingual task of MT to cross-lingual applications.

# Motivation

- Machine Translation (MT) has become increasingly commonplace in everyday use through BabelFish, Bing Translator or Google Translate.
- MT approaches provide good results but none of them exploit contextual information like named entity to help user comprehension.

## Solution

- Named **E**ntity **R**ecognition for the **I**nteractive **T**ranslation **S**ystem
- **The goal** is to propose a cross-lingual translation by providing detailed information extracted from DBpedia about persons, locations and organizations in the mother tongue of the user.
- This helps at scaling the traditional multilingual task of MT to cross-lingual applications.

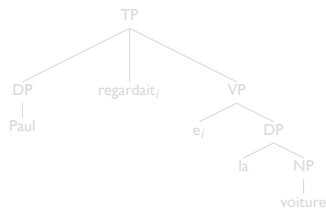
# ITS-2: an Interactive Translation System

Machine Translation system based on the multilingual parser Fips

## A classical syntactic transfer system:

1. Lexical and syntactic analysis with Fips
2. Lexical and grammatical transfer
3. Syntactic and morphological generation

Paul regardait la voiture → Paul was looking to the car



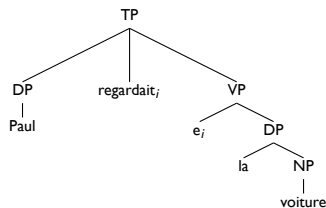
# ITS-2: an Interactive Translation System

Machine Translation system based on the multilingual parser Fips

## A classical syntactic transfer system:

1. Lexical and syntactic analysis with Fips
2. Lexical and grammatical transfer
3. Syntactic and morphological generation

Paul regardait la voiture → Paul was looking to the car



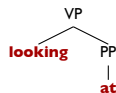
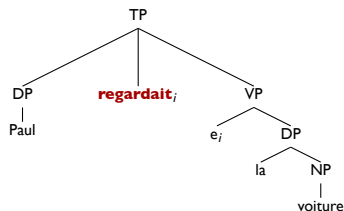
# ITS-2: an Interactive Translation System

Machine Translation system based on the multilingual parser Fips

## A classical syntactic transfer system:

1. Lexical and syntactic analysis with Fips
2. **Lexical and grammatical** transfer
3. Syntactic and morphological generation

**Paul regardait la voiture** → Paul was looking to the car



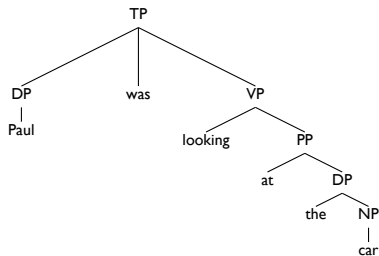
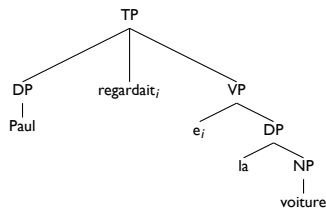
# ITS-2: an Interactive Translation System

Machine Translation system based on the multilingual parser Fips

## A classical syntactic transfer system:

1. Lexical and syntactic analysis with Fips
2. Lexical and grammatical transfer
3. Syntactic and morphological generation

Paul regardait la voiture → Paul was looking to the car



# Wikimeta: a Semantic Annotation platform

Wikimeta is a multilingual semantic annotation tool which performs:

- Named Entity Recognition (person, location, organization, etc.)
- Named Entity Linking

## Approach

- Contextual words resource called *Linked Data Interface* (LDI)
- Disambiguation task using *Semantic Disambiguation Algorithm* (SDA)



# Wikimeta: a Semantic Annotation platform

Wikimeta is a multilingual semantic annotation tool which performs:

- Named Entity Recognition (person, location, organization, etc.)
- Named Entity Linking

## Approach

- Contextual words resource called *Linked Data Interface* (LDI)
- Disambiguation task using *Semantic Disambiguation Algorithm* (SDA)

Washington D.C. (LOC) - Barack Obama (PERS)  
confronted a trio of controveuse, pledging to work with  
Congress (ORG) to ensure the IRS (ORG) doesn't abuse  
its power.

# Wikimeta: a Semantic Annotation platform

Wikimeta is a multilingual semantic annotation tool which performs:

- Named Entity Recognition (person, location, organization, etc.)
- Named Entity Linking

## Approach

- Contextual words resource called *Linked Data Interface* (LDI)
- Disambiguation task using *Semantic Disambiguation Algorithm* (SDA)

Washington D.C. (LOC) - Barack Obama (PERS)  
confronted a trio of controverse, pledging to work with  
Congress (ORG) to ensure the IRS (ORG) doesn't abuse  
its power.

[http://dbpedia.org/data/Washington\\_D.C..rdf](http://dbpedia.org/data/Washington_D.C..rdf)  
[http://dbpedia.org/data/Barack\\_Obama.rdf](http://dbpedia.org/data/Barack_Obama.rdf)  
[http://dbpedia.org/data/United\\_States\\_Congress.rdf](http://dbpedia.org/data/United_States_Congress.rdf)  
[http://dbpedia.org/data/Internal\\_Revenue\\_Service.rdf](http://dbpedia.org/data/Internal_Revenue_Service.rdf)

# Wikimeta: a Semantic Annotation platform

Wikimeta is a multilingual semantic annotation tool which performs:

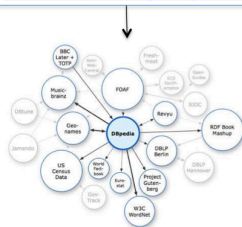
- Named Entity Recognition (person, location, organization, etc.)
- Named Entity Linking

## Approach

- Contextual words resource called *Linked Data Interface* (LDI)
- Disambiguation task using *Semantic Disambiguation Algorithm* (SDA)

Washington D.C. (LOC) - Barack Obama (PERS)  
confronted a trio of controveise, pledging to work with  
Congress (ORG) to ensure the IRS (ORG) doesn't abuse  
its power.

[http://dbpedia.org/data/Washington\\_D.C..rdf](http://dbpedia.org/data/Washington_D.C..rdf)  
[http://dbpedia.org/data/Barack\\_Obama.rdf](http://dbpedia.org/data/Barack_Obama.rdf)  
[http://dbpedia.org/data/United\\_States\\_Congress.rdf](http://dbpedia.org/data/United_States_Congress.rdf)  
[http://dbpedia.org/data/Internal\\_Revenue\\_Service.rdf](http://dbpedia.org/data/Internal_Revenue_Service.rdf)



# The NERITS Approach

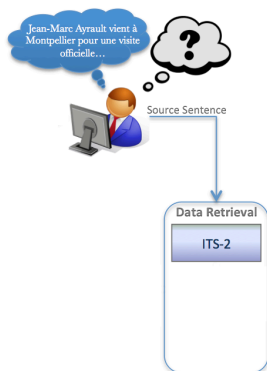
An example through NERITS Architecture



# The NERITS Approach

## Data Retrieval: Translation

<http://latlapps.unige.ch/Translation>

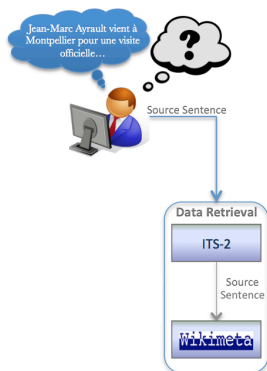


```
<Translation>
<input src="fr">
  Jean-Marc Ayrault vient à Montpellier pour
  une visite officielle.
</input>
<output tgt="en">
  Jean-Marc Ayrault is coming to
  Montpellier for an official visit.
</output>
</Translation>
```

# The NERITS Approach

## Data Retrieval: Semantic Annotation

<http://www.wikimeta.com/>

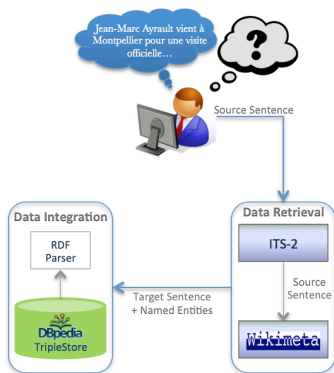


```
<extraction>
  <NE>Jean-Marc Ayrault</NE>
  <type>PERS.HUM</type>
  <LINKEDDATA>
    http://www.dbpedia.org/resource/Jean-Marc\_Ayrault
  </LINKEDDATA>
</extraction>
<extraction>
  <NE>Montpellier</NE>
  <type>LOC.ADMI</type>
  <LINKEDDATA>
    http://www.dbpedia.org/resource/Montpellier
  </LINKEDDATA>
</extraction>
```

# The NERITS Approach

## Data Integration

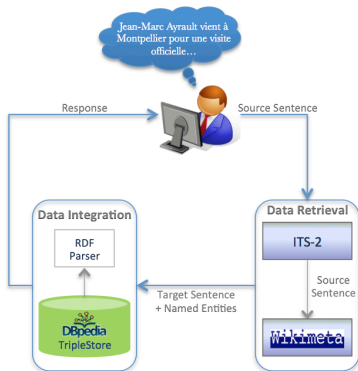
[http://dbpedia.org/data/Jean-Marc\\_Ayrault.rdf](http://dbpedia.org/data/Jean-Marc_Ayrault.rdf)




```
<rdf>
  <rdfs:label xml:lang="fr">Jean-Marc
    Ayrault</rdfs:label>
  <dbpedia-owl:orderInOffice>Prime
    Minister of France</dbpedia-
    owl:orderInOffice>
  <dbpprop:birthDate>1950-01-25</
    dbpprop:birthDate>
  <dbpprop:placeOfBirth>Maulévrier,
    Maine-et-Loire
    France</dbpprop:placeOfBirth>
  <dbpprop:spouse>Brigitte
    Terrien</dbpprop:spouse>
  <foaf:depiction rdf:resource="http://
    upload.wikimedia.org/wikipedia/
    commons/b/be/Jean-
    Marc_Ayrault_-_mars_2012.jpg" />
</rdf>
```

# The NERITS Approach

## Response



Jean-Marc Ayrault is coming to Montpellier for an official visit.



**Jean-Marc Ayrault**  
Occupation: Prime Minister of France  
Date of Birth: 1950-01-25  
Place of Birth: Maulévrier, Maine-et-Loire France  
Spouse: Brigitte Terrien

Jean-Marc Ayrault (born 25 January 1950) is the Prime Minister of France, appointed on 16 May 2012. He was the Mayor of Nantes from 1989 to 2012, and led the French Socialist Party group in the National Assembly from 1997 to 2012.





W

**Named Entity**  
Jean-Marc Ayrault



# The NERITS Approach

## Other Examples

	<p><b>United Nations</b> <b>Organisation des Nations unies</b></p> <p><b>Established Date:</b> 1945-06-26</p> <p><b>Admin Center:</b> in New York City, New York</p> <p>The United Nations (abbreviated UN in English, and ONU in its other official languages), is an international organization whose stated aims are facilitating cooperation in international law, international security, economic development, social progress, human rights, and achievement of world peace. The UN was founded in 1945 after World War II to replace the League of Nations, to stop wars between countries, and to provide a platform for dialogue.</p> <p><b>W</b></p>
	<p><b>Montpellier</b></p> <p><b>Population:</b> 252998</p> <p><b>Postal Code:</b> 34000</p> <p><b>Mayor:</b> Hélène Mandroux-Colas</p> <p>Montpellier is a city in southern France. It is the capital of the Languedoc-Roussillon region, as well as the Hérault department. Montpellier is the 8th largest city of the country, and is also the fastest growing city in France over the past 25 years.</p> <p><b>W</b></p>
	<p><b>Barack Obama</b></p> <p><b>Occupation:</b> President of the United States, 44th,</p> <p><b>Date of Birth:</b> 1961-08-04</p> <p><b>Place of Birth:</b> Honolulu, Hawaii, United States</p> <p><b>Spouse:</b> Michelle Robinson</p> <p>Barack Hussein Obama II is the 44th and current President of the United States. He is the first African American to hold the office. In January 2005, Obama was sworn in as a U.S. Senator in the state of Illinois. He would hold this office until November 2008, when he resigned following his victory in the 2008 presidential election. Born in Honolulu, Hawaii, Obama is a graduate of Columbia University and Harvard Law School, where he was the president of the Harvard Law Review.</p> <p><b>W</b> <b>the guardian</b> </p>

# The NERITS Approach

## User Interface

<http://cms.unige.ch/lettres/linguistique/nebhi/nerits/>

The screenshot displays the NERITS web application interface. On the left, a sidebar contains a logo and a descriptive paragraph about the system, along with the University of Geneva logo and contact information. The main area features a language selection dropdown set to 'French-English', an input box with the French sentence 'Jean-Marc Ayraud vient à Montpellier pour une visite officielle.', and buttons for 'Translate' and 'Reset'. Below the input, the translated English sentence is shown: 'Jean-Marc Ayraud is coming to Montpellier for an official visit.' To the right, a 'Legend' box identifies 'Location' (red), 'Organization' (orange), and 'Person' (blue). A 'Named Entity' box below it shows 'Jean-Marc Ayraud' as a person and 'Montpellier' as a location.

NERITS - A machine translation mashup system using Wikimeta and LOD

Machine Translation (MT) has become increasingly commonplace in everyday use through Web services such as Google Translate. Although the different MT approaches provide good results, none of them exploit contextual information like named entity - such as the names of persons, organizations or locations - to help user comprehension. In this application, we present a machine translation system using semantic annotation provided by Wikimeta and Linked Open Data from DBpedia. The goal of NERITS is to propose a cross-lingual translation by providing a set of information about persons, locations and organizations in the mother tongue of the user.

UNIVERSITÉ DE GENÈVE  
LABORATORY OF ANALYSIS AND TECHNOLOGY OF LANGUAGE

© 2013 LATL | Contact | @NERITS

French-English

Jean-Marc Ayraud vient à Montpellier pour une visite officielle.

Translate Reset

Jean-Marc Ayraud is coming to Montpellier for an official visit.

**Legend**

- Location
- Organization
- Person

**Named Entity**

- Jean-Marc Ayraud
- Montpellier

## Conclusion - Further Work

- We have presented a MT mashup system using Semantic Annotation provided by Wikimeta and LOD.
- We have given a first glance on how Semantic Annotation and information from DBpedia can be combined to help user comprehension.

### Further Work

- Recognition of Events (September 11 attacks, Fall of the Berlin Wall, etc.)
- Incorporate more knowledge from LOD (points of interest of a Location, people around a Person)
- Add other languages (French-German, German-French)

Questions?  
kamel.nebhi@unige.ch



Thank you!  
kamel.nebhi@unige.ch

