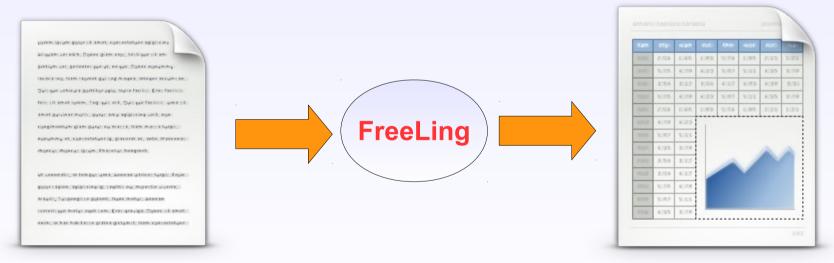
# FreeLing & TextServer: Cloud-based NLP

Lluís Padró
Centre de Recerca TALP
Universitat Politècnica de Catalunya
padro@cs.upc.edu



## What is FreeLing

- A configurable, extensible, developeroriented, linguistic analysis library.
- Linguistic Analysis: Obtain structured data from unstructured NL text.



- Similar stuff: Apache OpenNLP, NLTK
- Also related: UIMA, GATE

## FreeLing Background

#### Original goals:

 Provide a fast, reliable, and extensible platform to enable further progress in NLP application research and development.

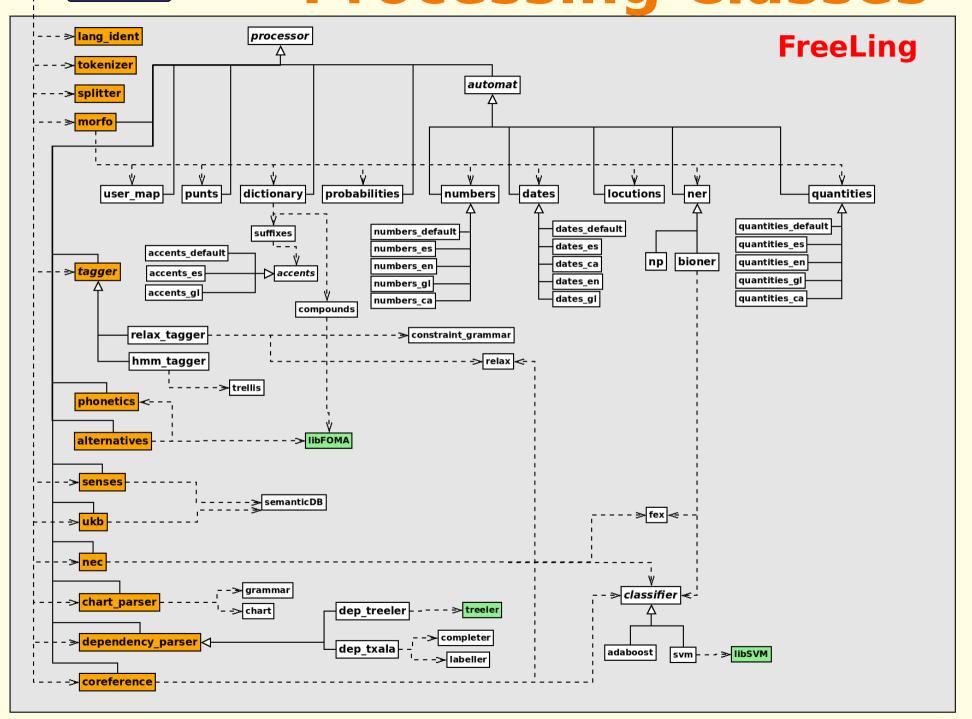
- 12+ year lasting open-source project
- 14 languages supported, thanks to contributors
- Continued improvement and extension
- Total ~200,000 downloads

## Functionalities & Languages

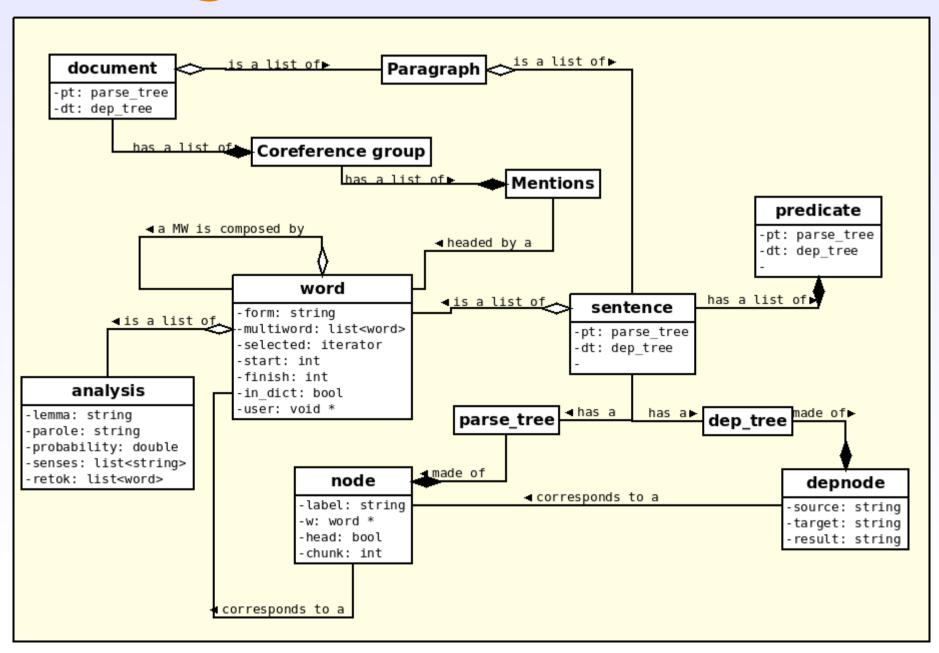
	as	ca	су	de	en	es	fr	gl	hr	it	nb	pt	ru	sl
Tokenization	X	X	X	X	X	X	X	X	Χ	X	Χ	X	Χ	X
Sentence splitting	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Number detection		X			X	X		X		X		X	X	
Date detection		X			X	X		X				X	X	
Morphological dictionary	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Affix rules	X	X	X	X	X	X	X	X		X	X	X		
Multiword detection	X	X	X		X	X	X	X		X		X		
Basic named entityd etection	X	X	X		X	X	X	X		X		X	X	X
B-I-O named entity detection		X			X	X		X				X		
Named Entity Classification		X			X	X						X		
Quantity detection		X			X	X		X				X	X	
PoS tagging	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Phonetic encoding					X	X								
WN sense annotation		X			X	X	X	X						X
UKB sense disambiguation		X			X	X	X							X
Shallow parsing	X	X			X	X		X				X		
Full/dependency parsing	X	X			X	X		X	X					Χ
Semantic Role Labelling		X			X	X			X					X
Coreference resolution					X	X								

Client Application

## **Processing Classes**



## Linguistic Data Classes



### **Demo**

http://nlp.lsi.upc.edu/freeling/demo/demo.php

#### **TextServer**

New requirements from application developers

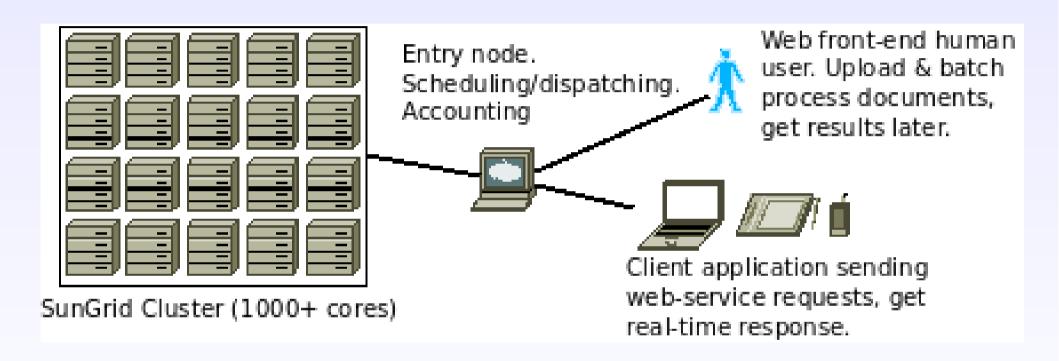
- Mobile devices
- Big data



Toolkits and libraries are moving into the cloud.

- Simplicity of installation, configuration, and use.
- Possibility of providing a wider and more flexible range of services.
- Ubiquity

### **TextServer Architecture**



TextServer currently provides FreeLing-based analyzers, but it aims to provide many other kinds of NLP related services.

## **TextServer Advantadges**

- Research
  - Replicability.
  - Reusability.
  - Valuable asset for UPC as project partner.
- Application development
  - Simplicity of use.
  - Speed, parallelism.
- Technology transfer
  - Less licensing problems.
  - Monetization of research results.

#### Demo

http://textserver.lsi.upc.edu/textserver

# FreeLing & TextServer: Cloud-based NLP

Lluís Padró
Centre de Recerca TALP
Universitat Politècnica de Catalunya
padro@cs.upc.edu

