Poetry Lab

POSTDATA Poetry Standardization and Linked Open Data

erc European Research Council — Established by the European Commission — Supporting top researchers from anywhere in the world

Elena González-Blanco| Salvador Ros Muñoz| Mª Luisa Díez Platas | Pablo Ruiz Fabo | Helena Bermúdez Sabel | Luciana Ayciriex | Clara Martínez Cantón

MAIN GOALS

a) Develop Tools for automatic poetry analysis, largely based on Natural Language Processing.
b) Carry out the detection of literary phenomena relied on linguistic characteristics.

What has it been already done?

- Content analysis (NERC)
  Medieval Spanish Named Entity Recognition)

- Stylistic traits detection
  Enjambment detection

ANJA
Automatic eNJambment Analysis in Spanish

- Detection of mismatch between metrical and syntactic units
- Defined in syntactic terms
- Detection based on pos-tagging / parsing
- Previously unanalyzed with automatic means

System architecture
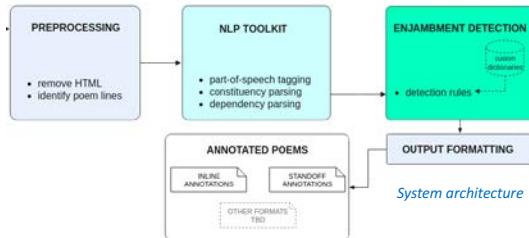
Output formats
- Standoff (type of enjambment and line)
- Inline (grammar tag and type of enjambment per line).

NLP Annotations
- PosTags tab (grammar tags)
- FullNLP tab (full pipeline)

Results
- Two different corpus: ss. XV-XIX and s. XX

		N	P	R	**F1**
20th	untyped	277	84.19	88.45	**86.27**
	typed	277	78.50	83.03	**80.70**

|		N	P	R	**F1**
Dia-chronic	untyped	260	74.18	87.64	**80.35**
	typed	260	61.24	72.31	**66.31**

HisMeTag
Hispanic Medieval Tagging

- Named Entities Recognition in Medieval Spanish texts
- NLP techniques for linguistic processing
- Normalization of historical spelling and variant generation
- Automatic Detected Names Tagging and New proposed names

System architecture

Dependence analysis

don Alfonso por la gracia de Dios rey de Castiella de Leon …
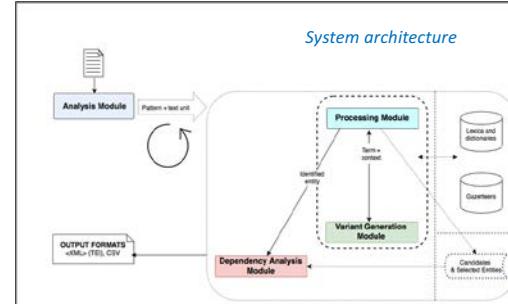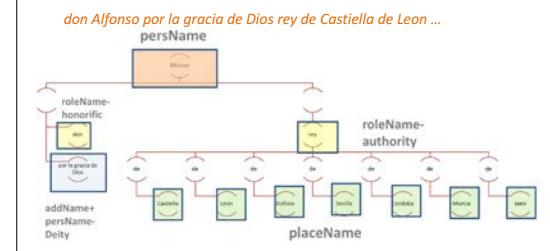
Output Formats

- TEI tagged Ouput

- Geolocation

Results

Entity type	Precision	Recall	F1
overall	0.83	0.72	0.77
placeName	0.78	0.73	0.75
persName	0.84	0.79	0.81
orgName	0.93	0.81	0.87
roleName	0.87	0.65	0.74

Period (years)	F1
1150-1200	0.85
1200-1250	0.73
1250-1300	0.89
1300-1350	0.76
1350-1370	0.63
1400-1500	0.67

http://prototipo-postdata.linhd.es/poetry-lab.php

UNED

LiNHD LABORATORIO DE INNOVACIÓN EN HUMANIDADES DIGITALES

Scan me