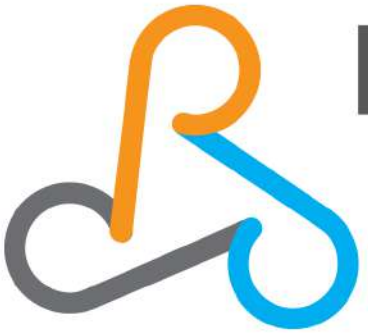


POSTDATA

Poetry Standardization and Linked Open Data

Humanidades digitales, tecnologías del lenguaje y análisis de poesía.

Elena González-Blanco García
UNED - Coverwallet
elena@coverwallet.com

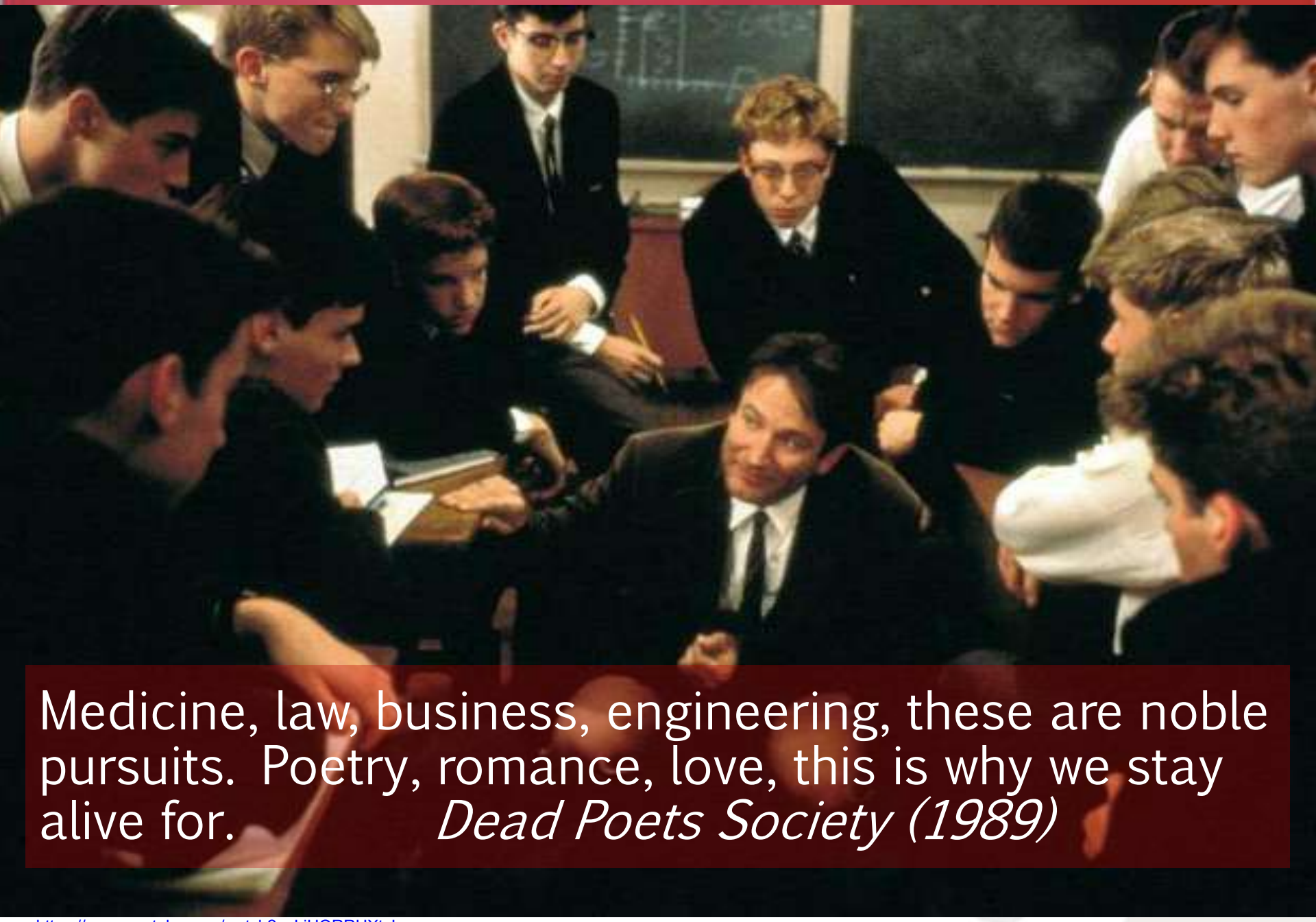
UNED

LINHDI LABORATORIO DE INNOVACIÓN EN HUMANIDADES DIGITALES

erc European Research Council
Supporting top researchers from anywhere in the world

¿Cómo analizamos la poesía?

POSTDATA
Poetry Standardization and Linked Open Data



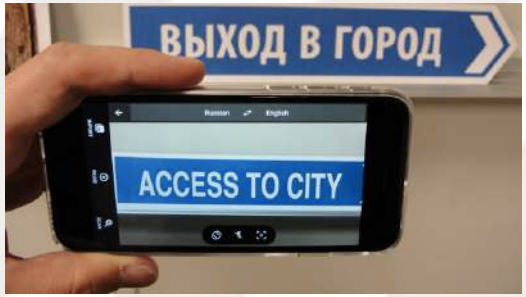
Medicine, law, business, engineering, these are noble pursuits. Poetry, romance, love, this is why we stay alive for. *Dead Poets Society (1989)*

https://www.youtube.com/watch?v=LjHORRHXtyI

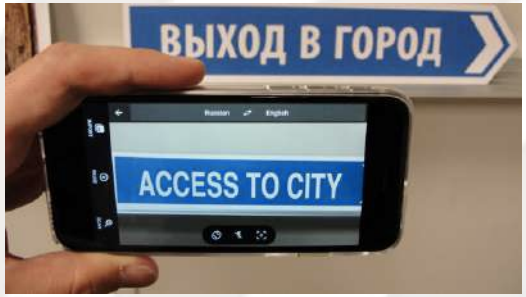
LAS TECNOLOGÍAS DEL LENGUAJE

LAS TECNOLOGÍAS DEL LENGUAJE

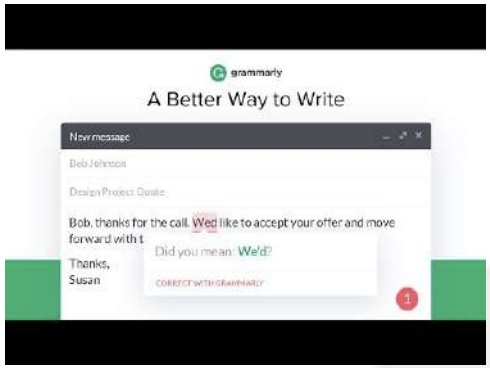
LAS TECNOLOGÍAS DEL LENGUAJE

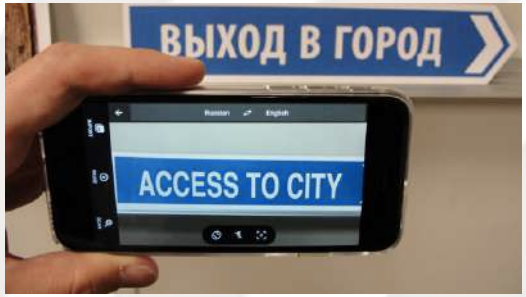


LAS TECNOLOGÍAS DEL LENGUAJE

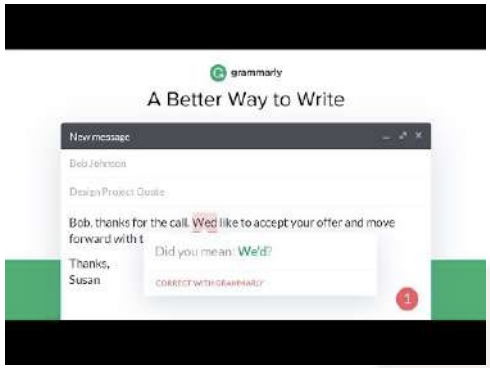


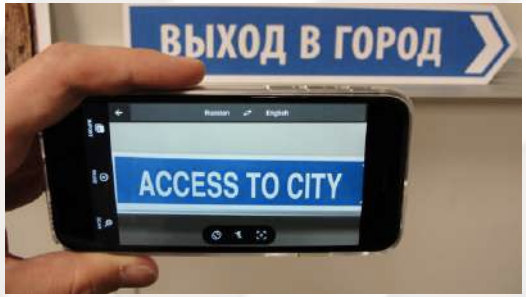
LAS TECNOLOGÍAS DEL LENGUAJE



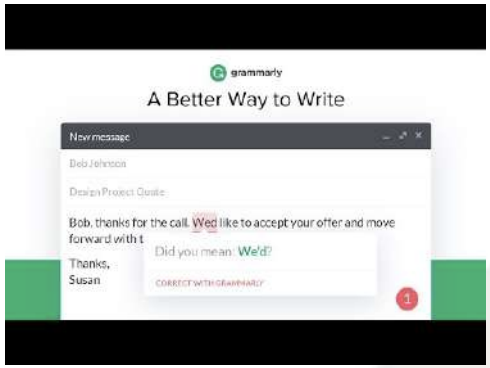


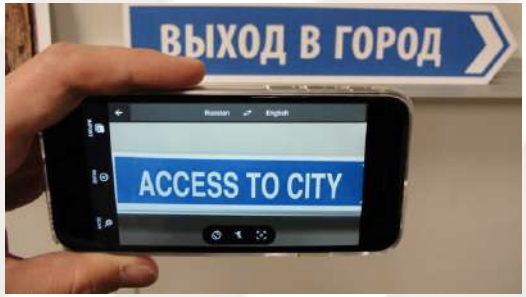
LAS TECNOLOGÍAS DEL LENGUAJE





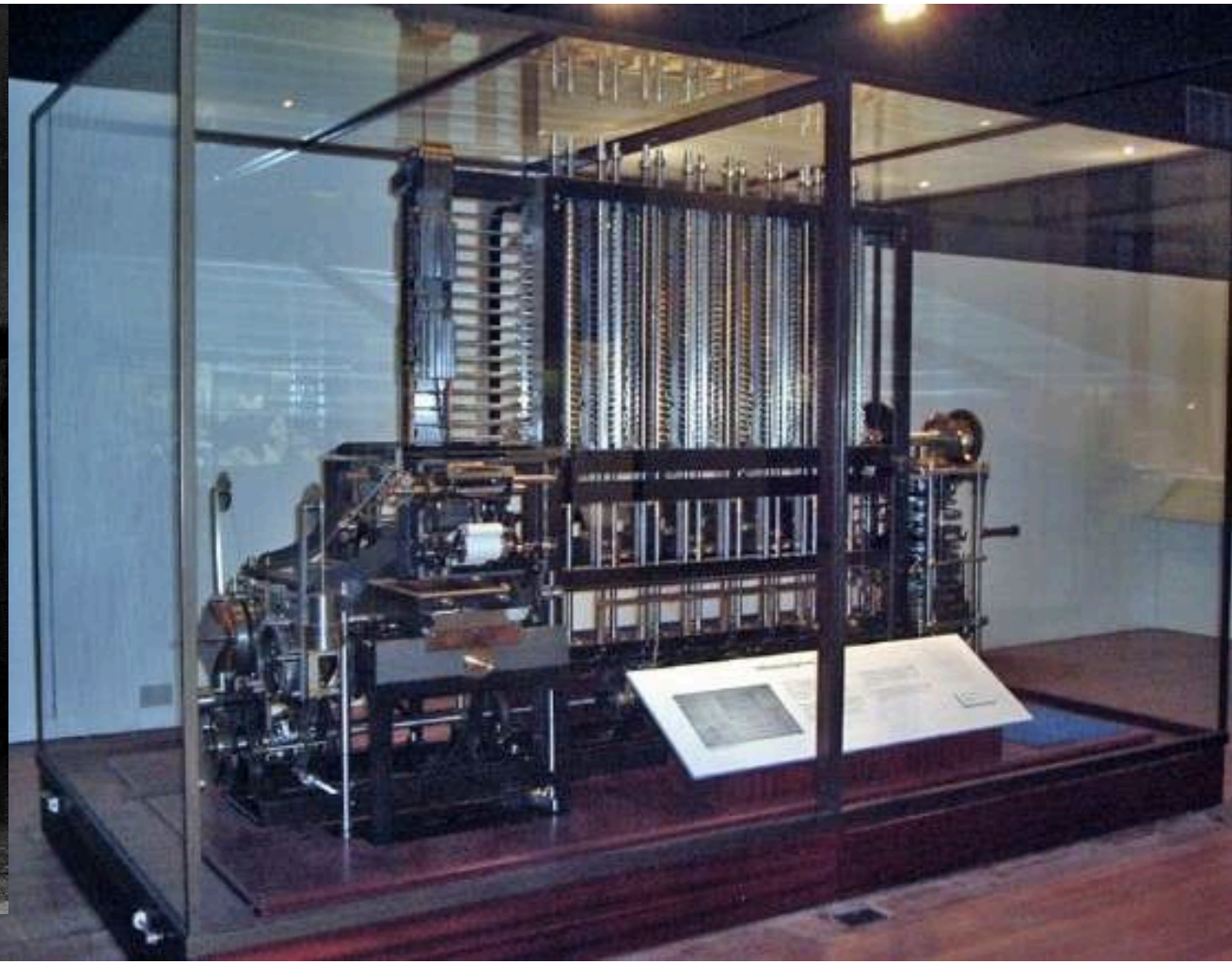
LAS TECNOLOGÍAS DEL LENGUAJE





ADA LOVELACE, 1815-1842

POSTDATA
Poetry Standardization and Linked Open Data



Alan Turing, 1912-54 “Turing Test”

VOL. LIX. No. 236.] [October, 1950

MIND

A QUARTERLY REVIEW

OF

PSYCHOLOGY AND PHILOSOPHY

I.—COMPUTING MACHINERY AND INTELLIGENCE

BY A. M. TURING

1. *The Imitation Game.*

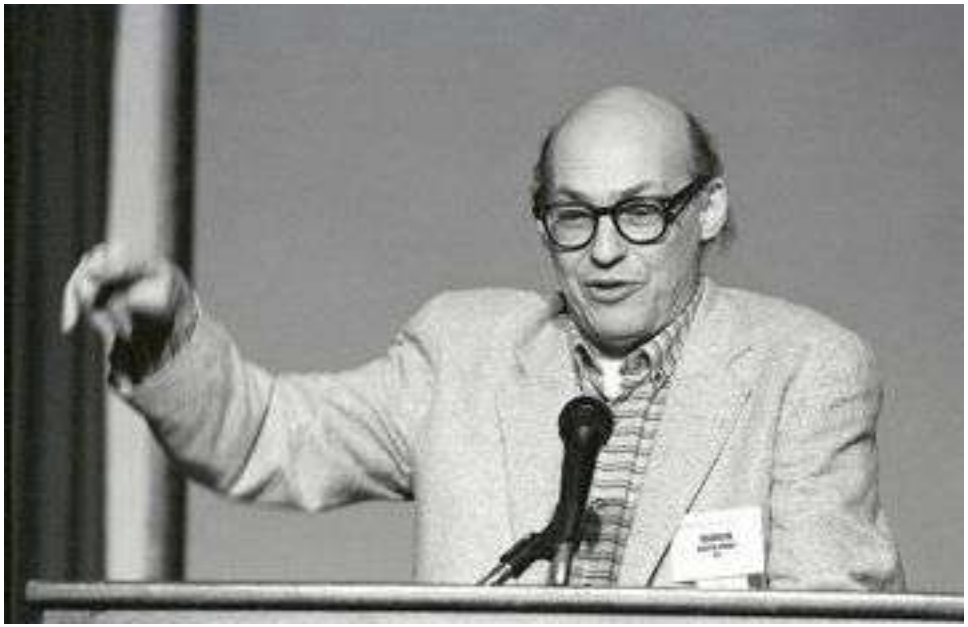
I PROPOSE to consider the question, 'Can machines think?' This should begin with definitions of the meaning of the terms 'machine' and 'think'. The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words 'machine' and 'think' are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, 'Can machines think?' is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.

The new form of the problem can be described in terms of a game which we call the 'imitation game'. It is played with three people, a man (A), a woman (B), and an interrogator (C) who may be of either sex. The interrogator stays in a room apart from the other two. The object of the game for the interrogator is to determine which of the other two is the man and which is the woman. He knows them by labels X and Y, and at the end of the game he says either 'X is A and Y is B' or 'X is B and Y is A'. The interrogator is allowed to put questions to A and B thus:

C: Will X please tell me the length of his or her hair?

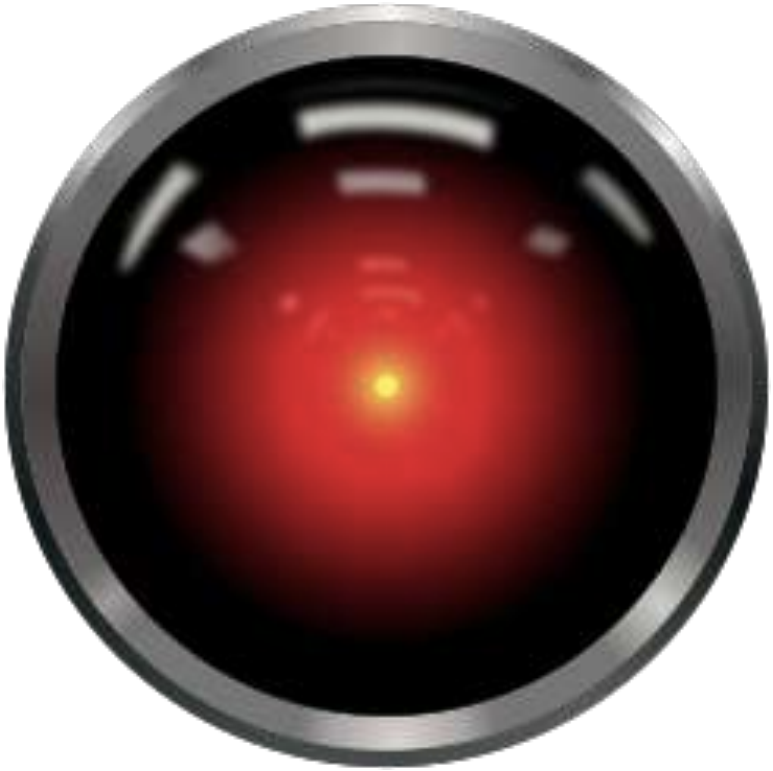
Now suppose X is actually A, then A must answer. It is A's

Marvin Minsky, 1927-2016



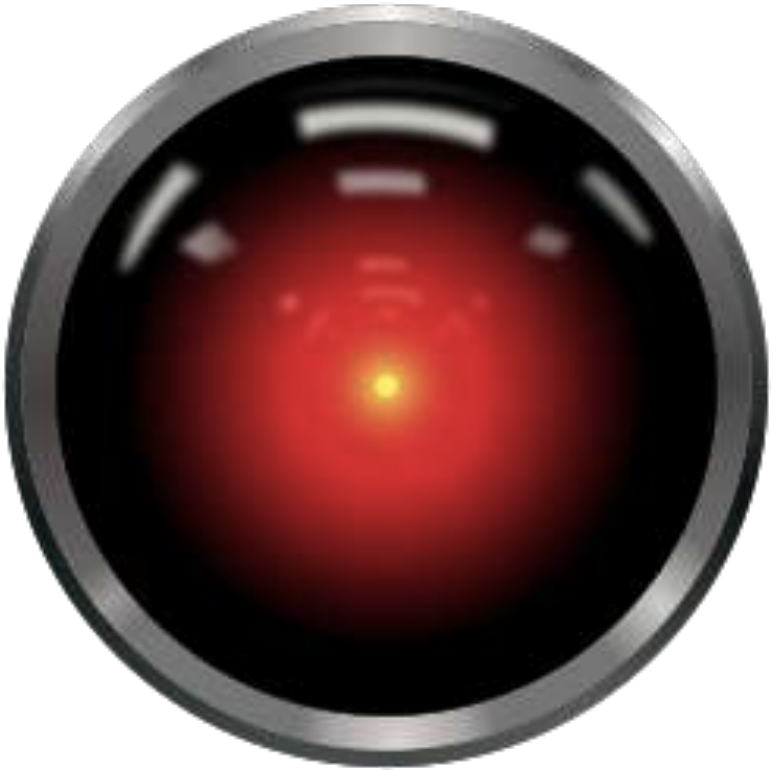
- Cofundador del Laboratorio de IA de MIT
- El “cerebro de la IA” – convencido de que las máquinas podían pensar como los hombres
- Investigación fundacional para la creación de redes neuronales artificiales.
- Asesor de *Odisea en el Espacio*, 2001, de Stanley Kubrick

clideo.com



2001 A Space Odyssey, 1968

POSTDATA
Poetry Standardization and Linked Open Data



2001 A Space Odyssey, 1968

LOS HITOS MÁS RELEVANTES

1950-60

Orígenes de la IA

Primeros algoritmos
Sistemas expertos
Fundamentos del NLP

1980-99

Estadística y modelos entrenados

Algoritmos probabilísticos
Entrenamiento con datos
Inclusión de gramáticas complejas y generación del lenguaje

2010-hoy

Redes neuronales, ML, DL

Redes neuronales y Deep Learning
Aprendizaje no supervisado
NLP combinado con DL - Neural Machine Translation

¿CÓMO APRENDEN LAS MÁQUINAS A HABLAR?

POSTDATA
Poetry Standardization and Linked Open Data

- Juancomiounamanzana

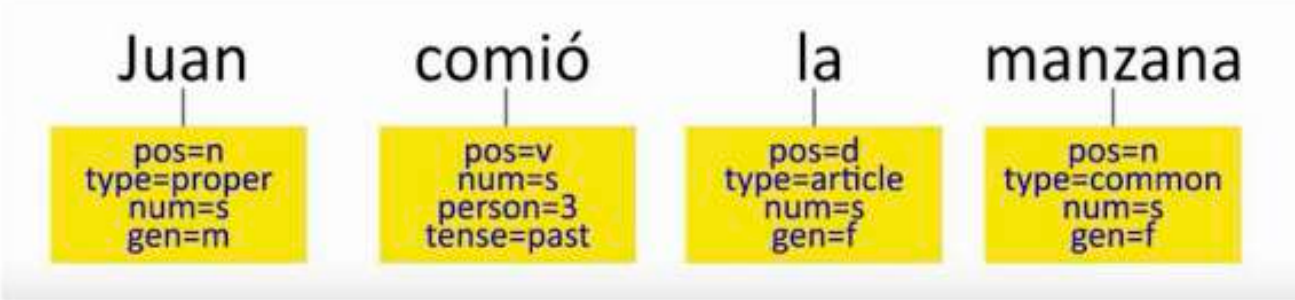
¿CÓMO APRENDEN LAS MÁQUINAS A HABLAR?

POSTDATA
Poetry Standardization and Linked Open Data

- Juan comio una manzana

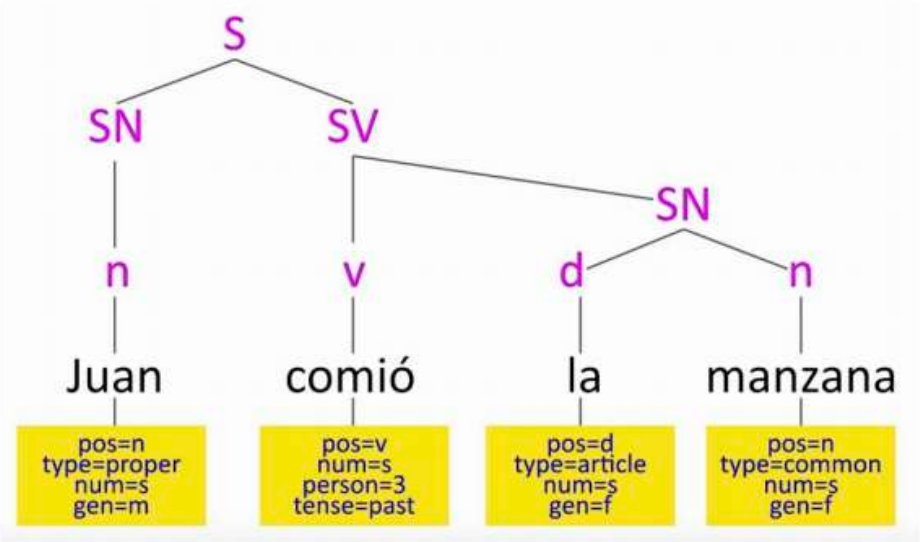
¿CÓMO APRENDEN LAS MÁQUINAS A HABLAR?

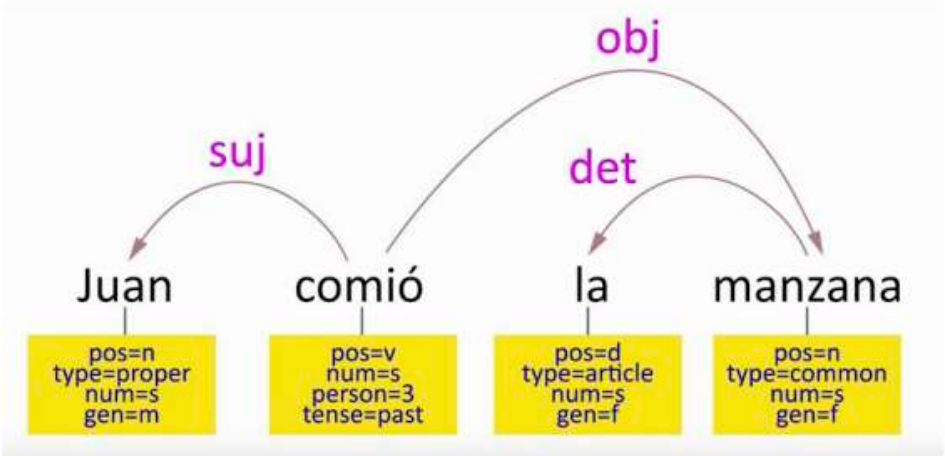
POSTDATA
Poetry Standardization and Linked Open Data



¿CÓMO APRENDEN LAS MÁQUINAS A HABLAR?

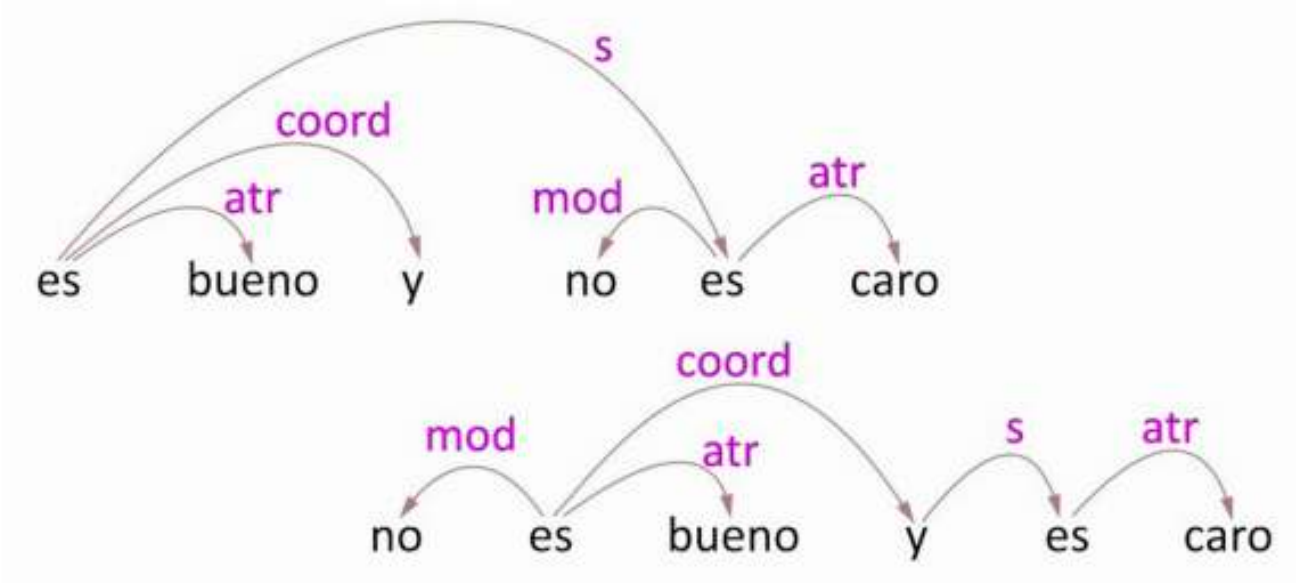
POSTDATA
Poetry Standardization and Linked Open Data



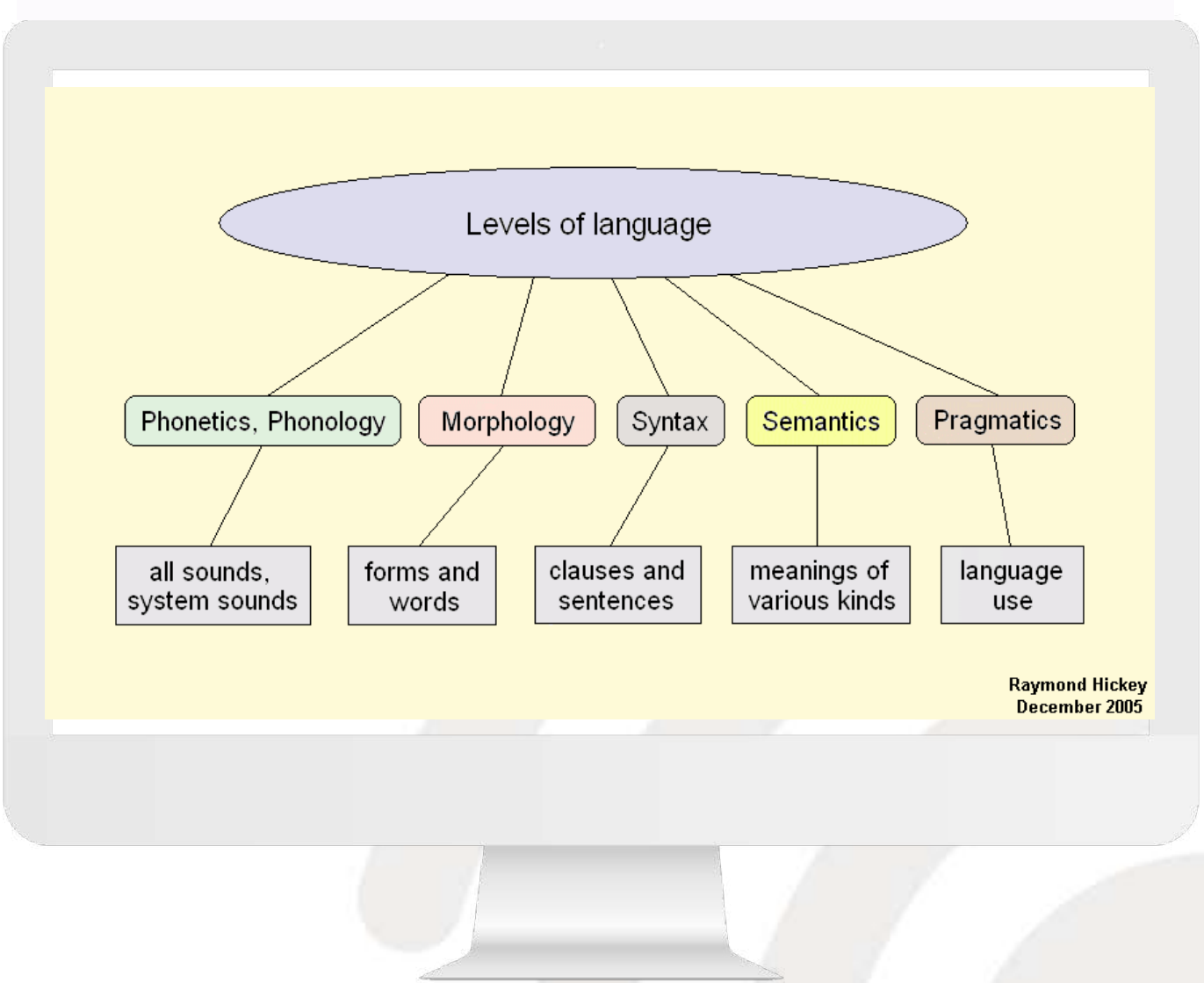


¿CÓMO APRENDEN LAS MÁQUINAS A HABLAR?

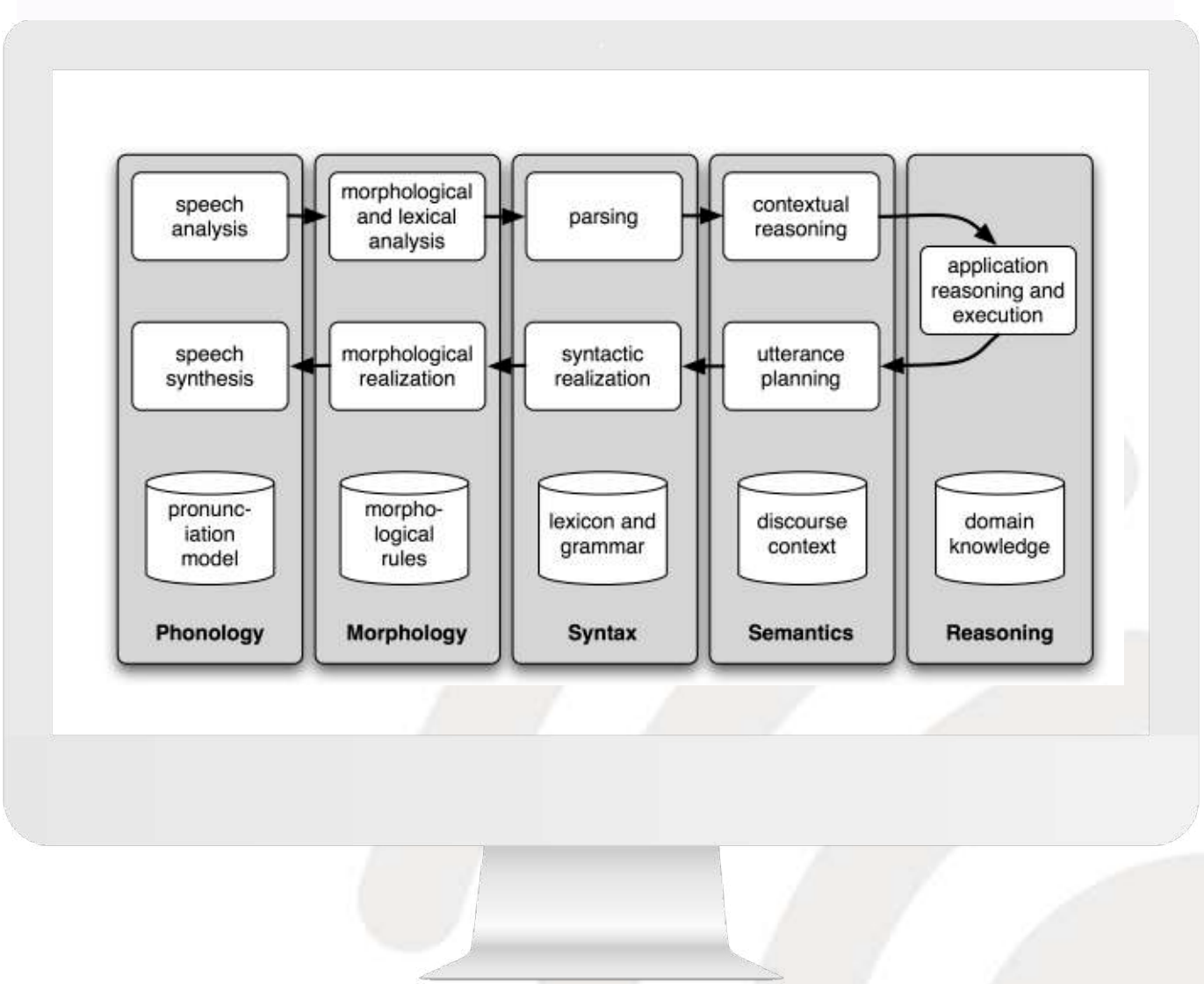
POSTDATA
Poetry Standardization and Linked Open Data



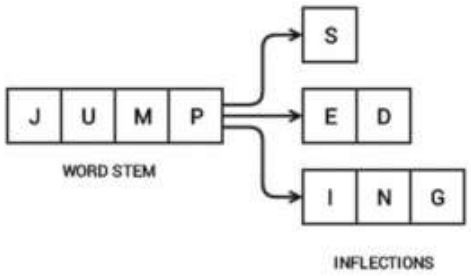
¿CÓMO ENSEÑAMOS TODO ESTO A UN ORDENADOR?



¿CÓMO ENSEÑAMOS TODO ESTO A UN ORDENADOR?



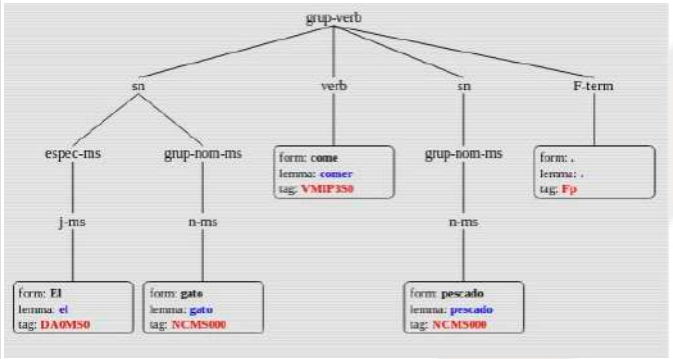
¿CÓMO ENSEÑAMOS TODO ESTO A UN ORDENADOR?



Part-Of-Speech Tagger - Análisis morfológico

Yo	bajo	con	el	hombre	bajo	a	tocar	el	bajo	bajo	la	escalera	.
yo	*bajar*	*con*	*el*	*hombre*	*bajo*	*a*	*tocar*	*el*	*bajo*	*bajo*	*el*	*escalera*	*.*
PP1CSN00	VMIP1S0	SPS00	DA0MS0	NCMS000	AQ0MS0	SPS00	VMN0000	DA0MS0	NCMS000	SPS00	DA0FS0	NCFS000	Fp

Treebank - Anotación sintáctica



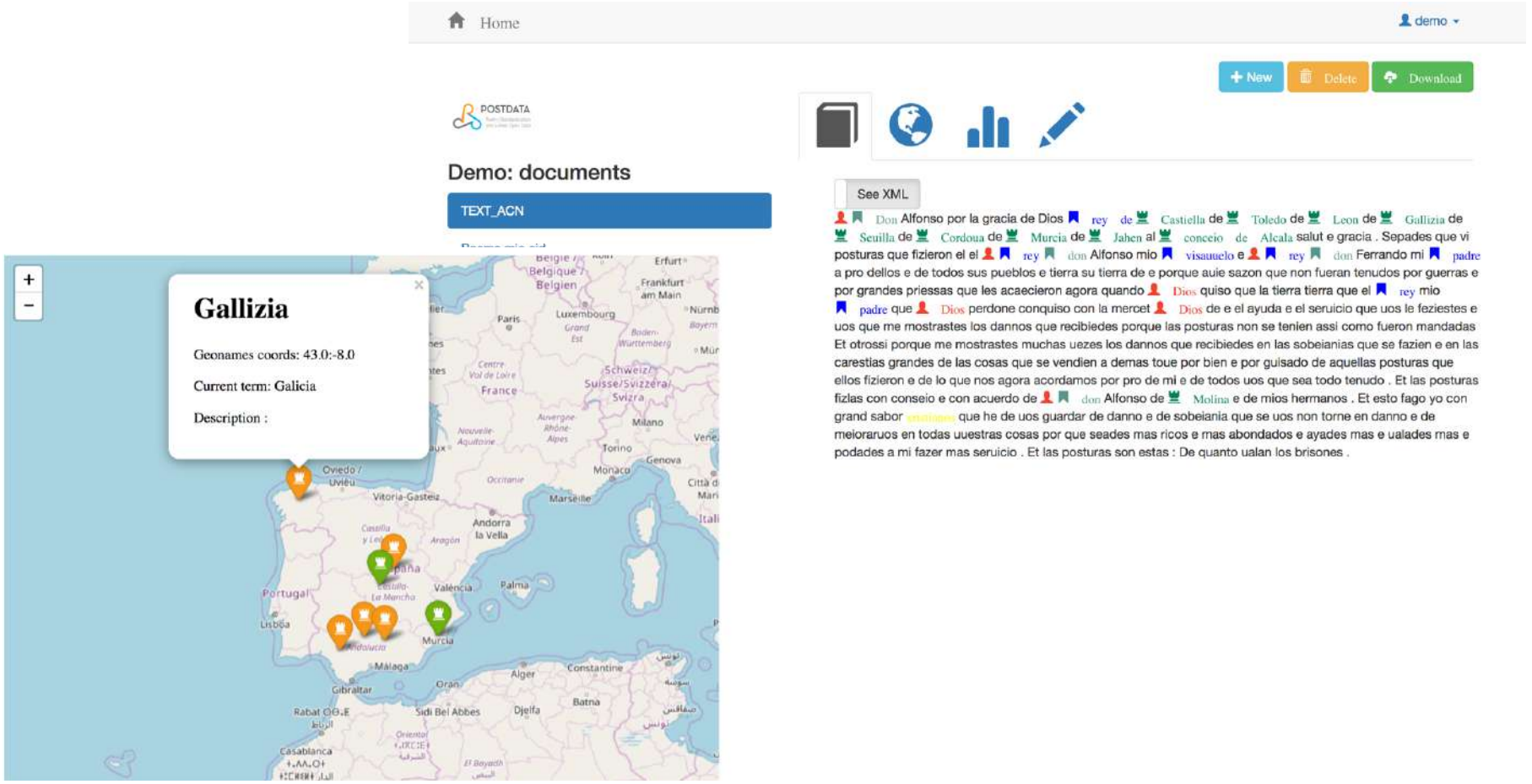
Wordnet - Análisis semántico

Necesito un gato para cambiar la rueda del coche.

Necesito necesitar VMIP1S0 01188725-v
un uno DI0MS0
gato gato NCMS000 03588951-n
para para SPS00
cambiar cambiar VMN0000 00126264-v
la el DA0FS0
rueda rueda NCFS000 04574999-n
de de SPS00
el el DA0MS0
coche coche NCMS000 02959942-n
. . Fp 1 -

NAME-ENTITY RECOGNITION - EXTRACCIÓN DE ENTIDADES

POSTDATA
Poetry Standardization and Linked Open Data



ANÁLISIS DE SENTIMIENTO

POSTDATA
Poetry Standardization and Linked Open Data

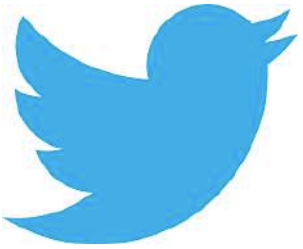
Dick's Sporting Goods

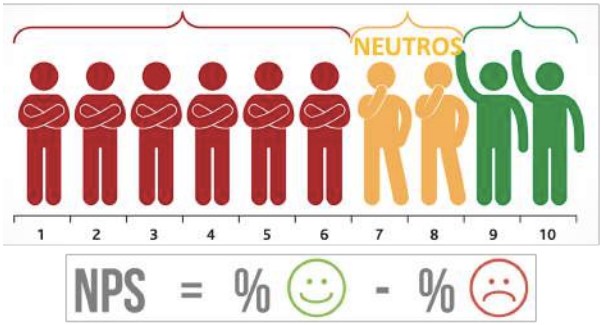
Seller rating: 4.4 / 5 - Based on 10,544 reviews

1 | 2 | 3 | 4 stars | 5 stars

What people are saying

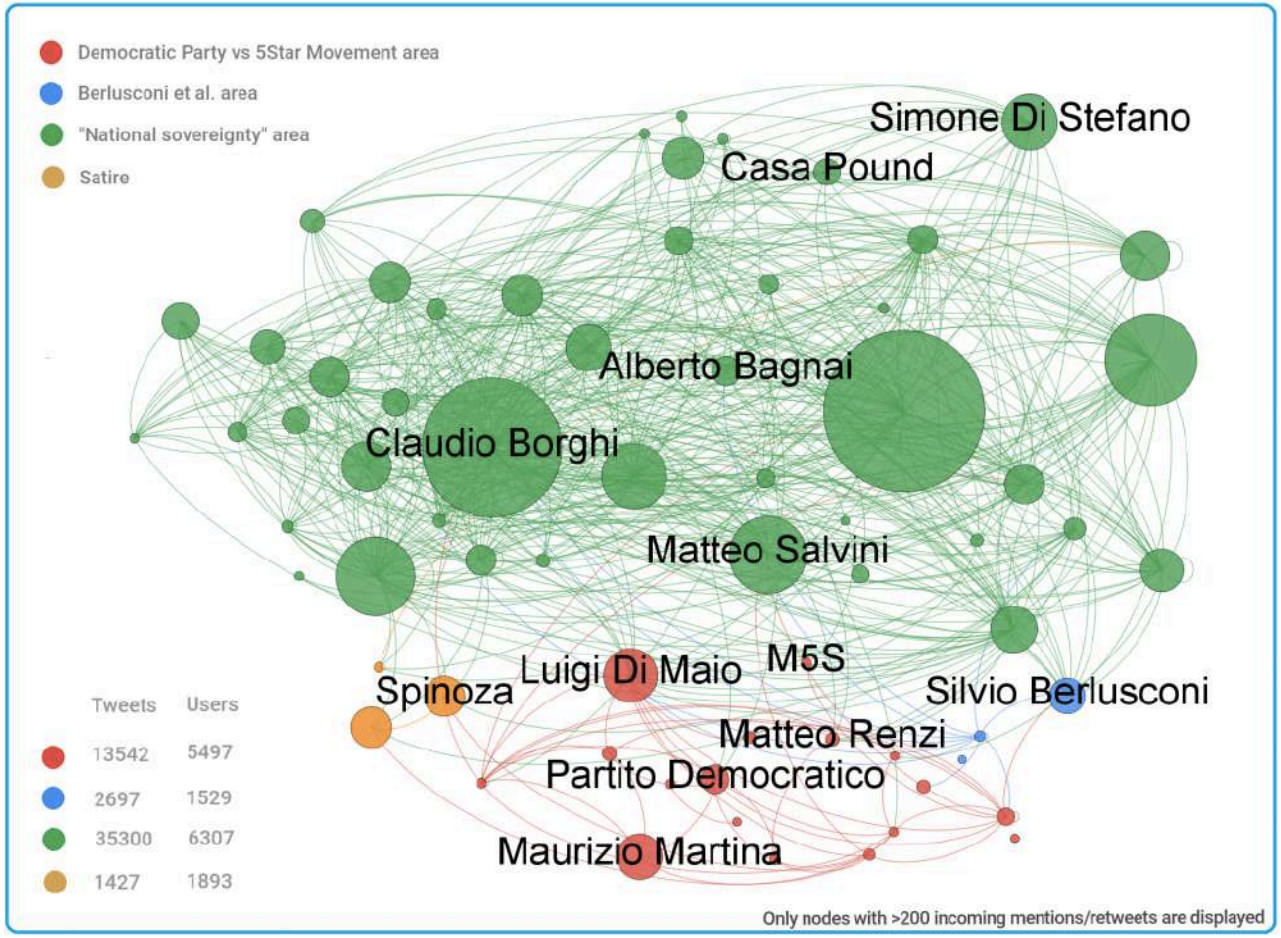
customer service	"Terrible customer service."
shipping	"Over all delivery speed was good."
price	"Great price, fast shipping, great product."
selection	"Fairly good selection of parts."
return policy	"Horrible return/exchange policy."
ordering process	"Really great transaction."
communication	"Quick shipping, great shipping communication"





ELECCIONES ITALIANAS 2018 - REPUTACIÓN DIGITAL

POSTDATA
Poetry Standardization and Linked Open Data



DETECCIÓN DE PLAGIO

VISUALIZING

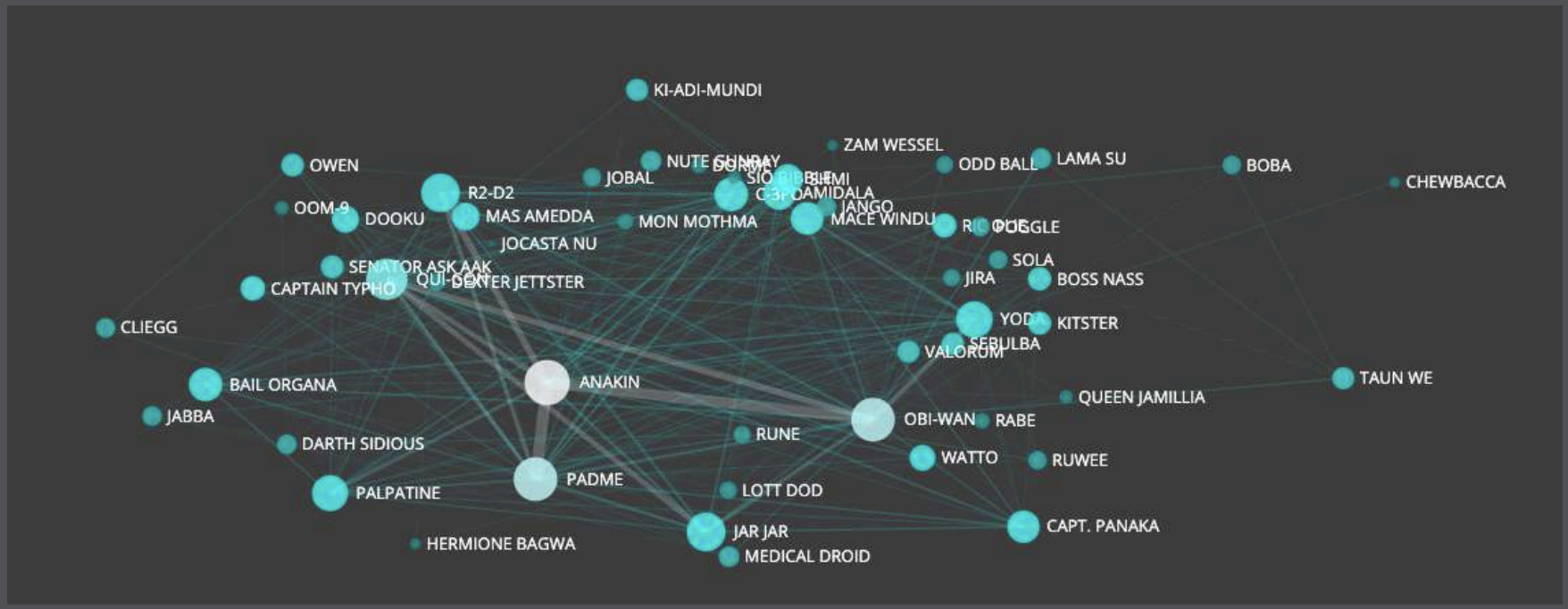
MOVIE SCRIPTS

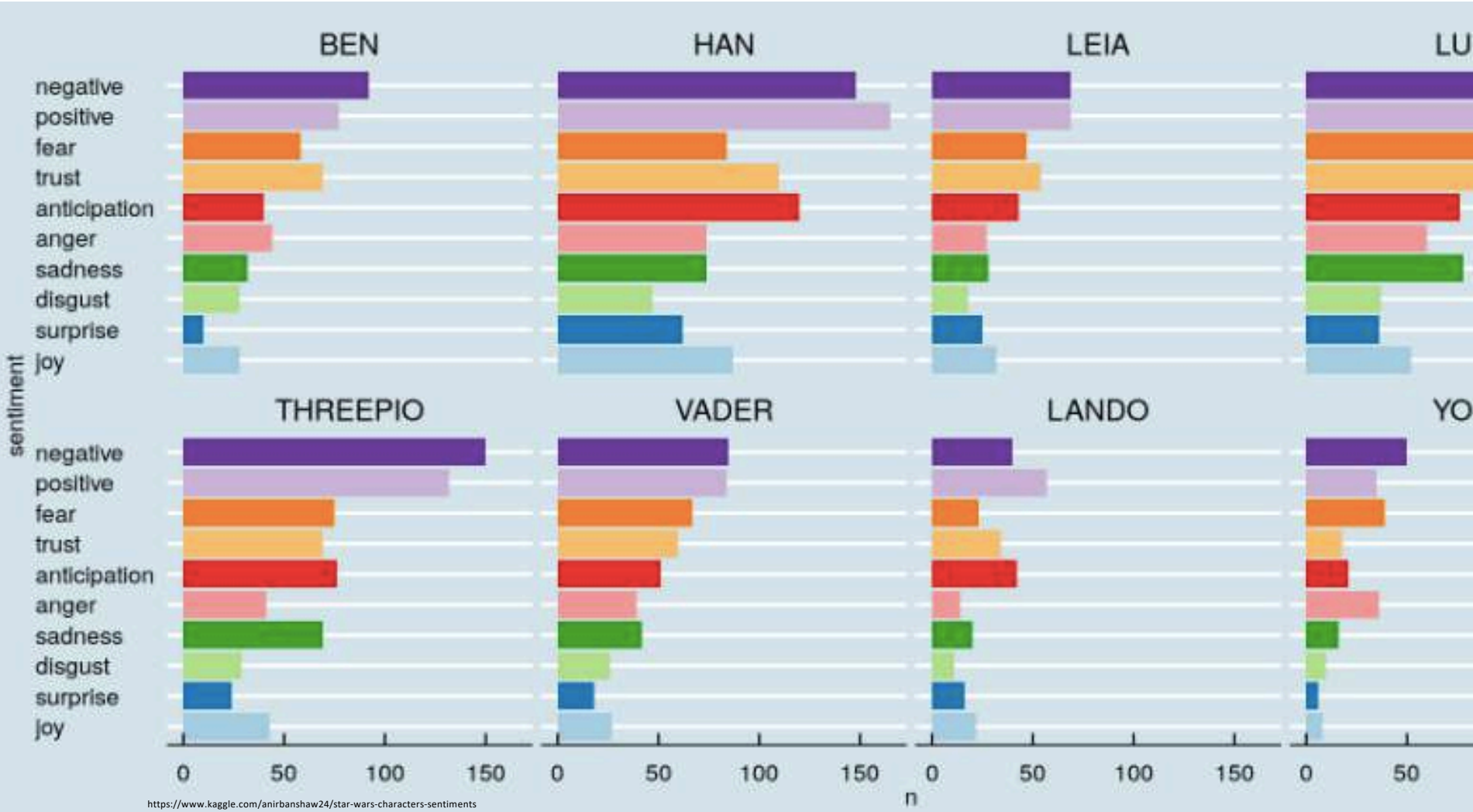
A long time ago, in a galaxy far, far away data analysts were talking about the upcoming new Star Wars movie. One of them has never seen any eposide before, so they decided to make the movie more accessible to this poor fellow.

http://labs.precognox.com/star-wars-visualization/

VISUALIZING

STAR WARS





https://www.kaggle.com/anirbanshaw24/star-wars-characters-sentiments

The Internet Movie Script Database (IMSDb)

IMSDb
The web's largest movie script resource!

DISFRUTA TU ESTANCIA
CON INFINITE INDULGENCE®.
TODO ESTÁ INCLUIDO.
DESCUBRE MÁS
GRAND PALLADIUM
HOTELS & RESORTS
PUNTA CANA

Search IMSDb

Go!

Alphabetical

#	A	B	C	D	E	F	G	H
I	J	K	L	M	N	O	P	Q
R	S	T	U	V	W	X	Y	Z

Genre

Action	Adventure	Animation
Comedy	Crime	Drama
Family	Fantasy	Film-Noir
Horror	Musical	Mystery
Romance	Sci-Fi	Short
Thriller	War	Western

Sponsor

Last added

Secret Life of Walter Mitty, The (2013)

Theory of Everything, The (2014)

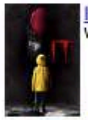
Hurt Locker, The (2009)

Thor Ragnarok (2017)

More >

Newest releases

Coco
Written by
Lee Unkrich
Jason Katz
Matthew Aldrich
Adrian Molina


It
Written by
Chase Palmer
Cary Fukunaga

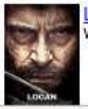
Big Sick, The
Written by
Emily V. Gordon
Kumail Nanjiani

Beauty and the Beast
Written by
Stephen Chbosky
Evan Spiliotopoulos

Thor Ragnarok
Written by
Eric Pearson
Craig Kyle
Christopher L. Yost

War for the Planet of the Apes
Written by
Mark Bomback
Matt Reeves

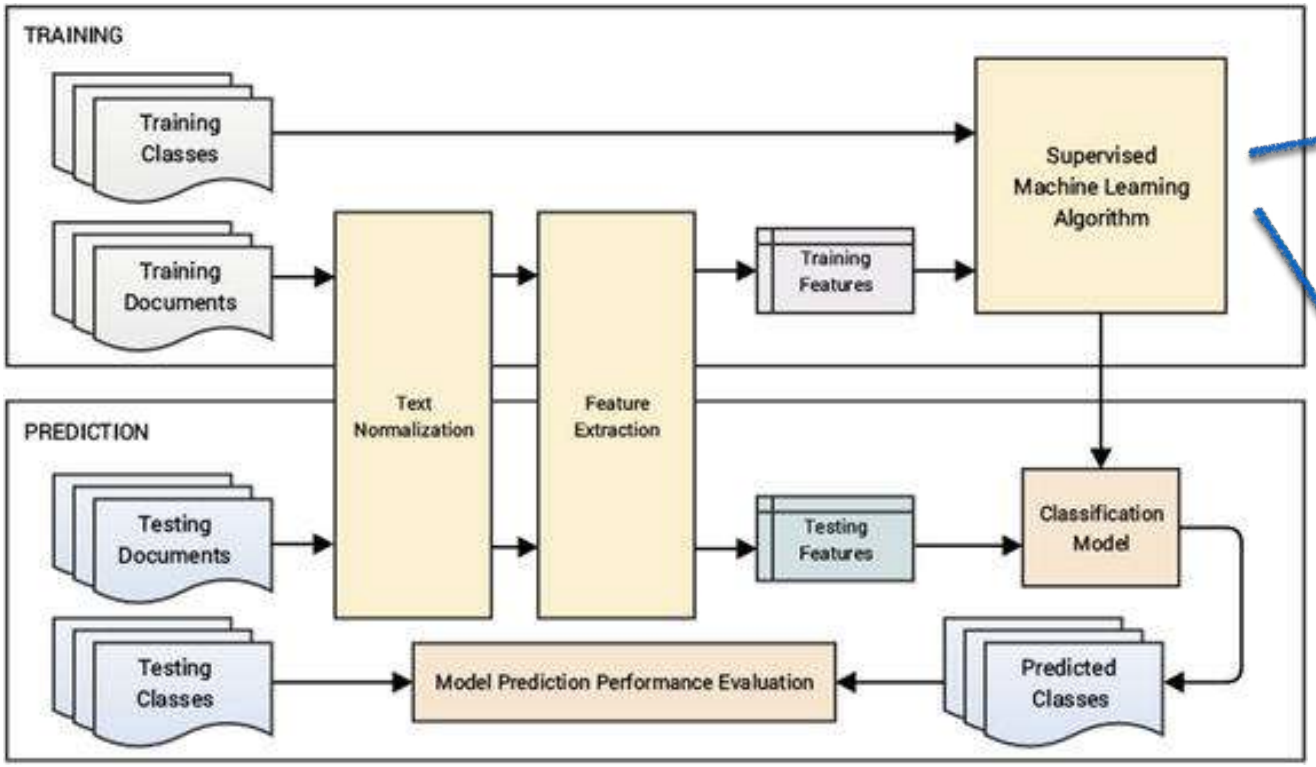
Guardians of the Galaxy Vol 2
Written by
James Gunn

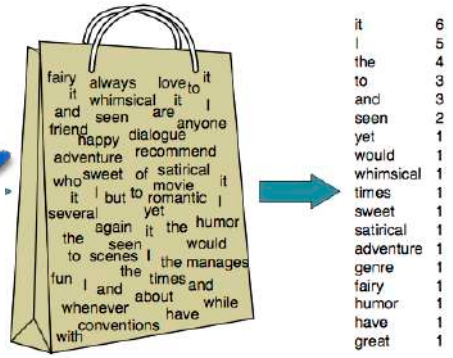

Logan
Written by
Scott Frank
James Mangold
Michael Green

https://www.imsdb.com/

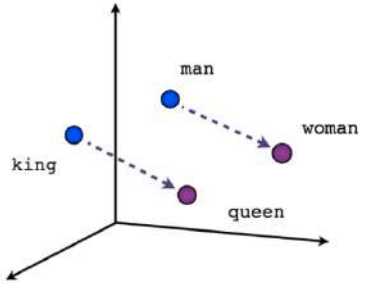
CLASIFICADORES DE TEXTOS

POSTDATA
Poetry Standardization and Linked Open Data





Naive Bayes



Male-Female

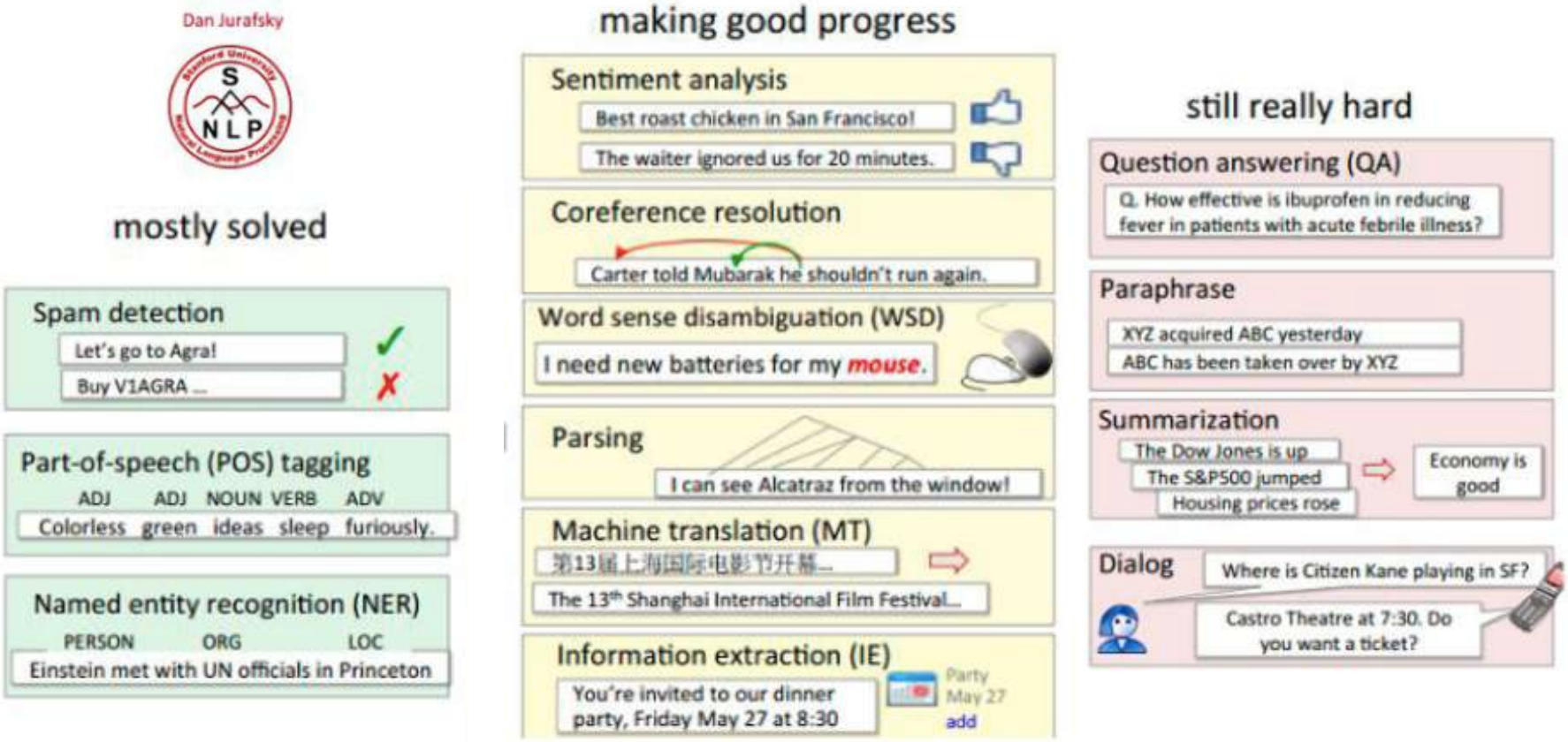
Word2Vec

¿Qué podemos hacer con tecnologías del lenguaje?

POSTDATA
Poetry Standardization and Linked Open Data

¿Qué podemos hacer con tecnologías del lenguaje?

POSTDATA
Poetry Standardization and Linked Open Data



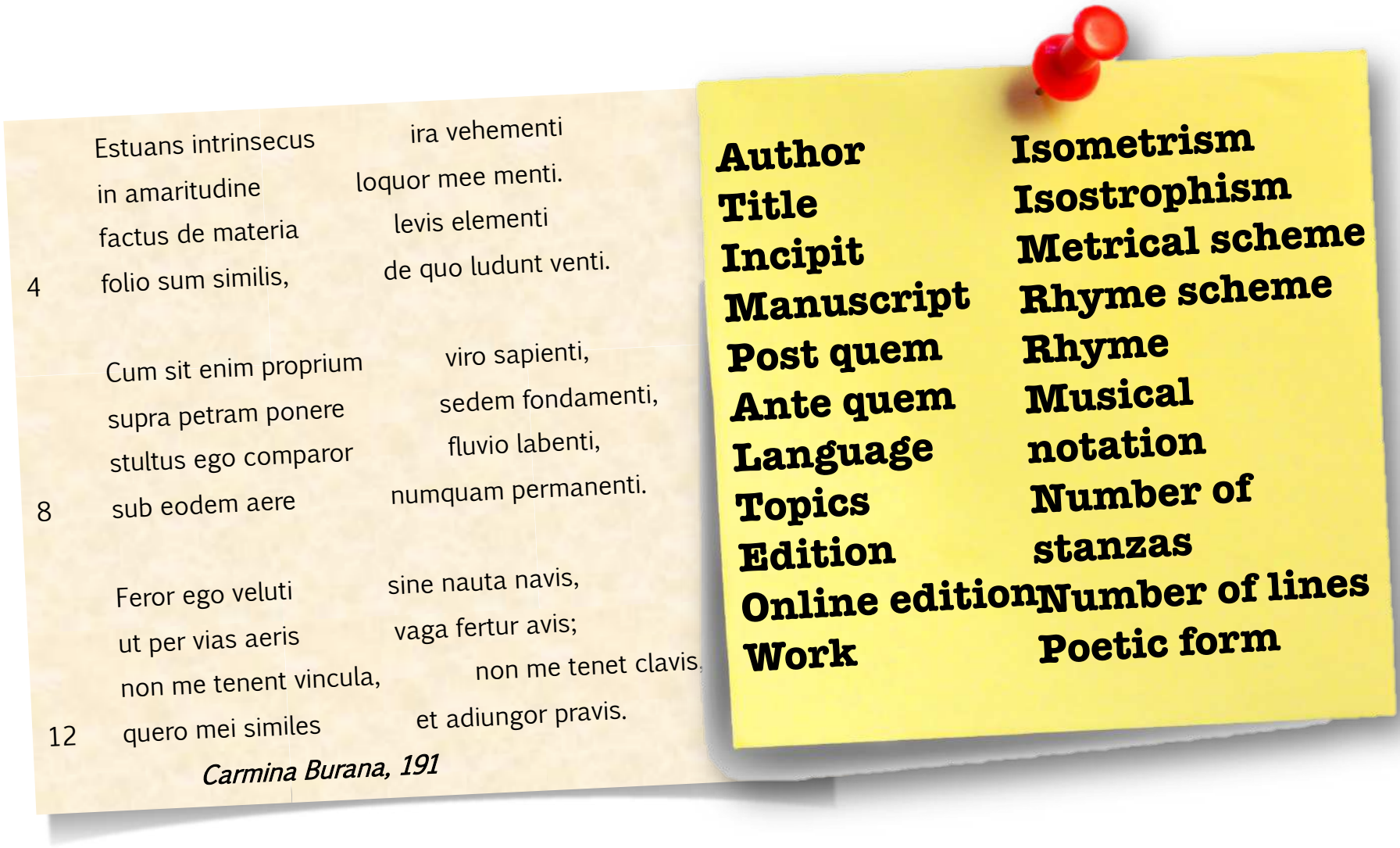
¿Cómo se concibe un proyecto de poesía desde las Humanidades Digitales?

POSTDATA
Poetry Standardization and Linked Open Data

Traditional Research		Digital Humanities
Counting syllables with fingers	→	Automatic syllable counter
Comparing 2-3 poems at the same time	→	Comparing thousands of poems in parallel
Looking for similar evidences by reading all papers and books	→	Finding multiple evidences through one search engine
Individual research	→	Collaborative, interdisciplinary research

¿Cómo definimos un poema?

POSTDATA
Poetry Standardization and Linked Open Data



Estuans intrinsecus ira vehementi
in amaritudine loquor mee menti.
factus de materia levis elementi
4 folio sum similis, de quo ludunt venti.

Cum sit enim proprium viro sapienti,
supra petram ponere sedem fondamenti,
stultus ego comparor fluvio labenti,
8 sub eodem aere numquam permanenti.

Feror ego veluti sine nauta navis,
ut per vias aeris vaga fertur avis;
non me tenent vincula, non me tenet clavis,
12 quero mei similes et adiungor pravis.

Carmina Burana, 191

Author
Title
Incipit
Manuscript
Post quem
Ante quem
Language
Topics
Edition
Online edition
Work
Isometrism
Isostrophism
Metrical scheme
Rhyme scheme
Rhyme
Musical notation
Number of stanzas
Number of lines
Poetic form

Barreras filológicas

POSTDATA
Poetry Standardization and Linked Open Data

Estuans intrinsecus ira vehementi
in amaritudine loquor mee menti.
factus de materia levis elementi
4 folio sum similis, de quo ludunt venti.

Cum sit enim proprium viro sapienti,
supra petram ponere sedem fondamenti,
stultus ego comparor fluvio labenti,
8 sub eodem aere numquam permanenti.

Feror ego veluti sine nauta navis,
ut per vias aeris vaga fertur avis;
non me tenent vincula, non me tenet clavis,
12 quero mei similes et adiungor pravis.

Carmina Burana, 191

Alexandrines

Goliardic

12A12A12A12A
(Romance)

4x(7pp+7p)
(Classic Latin)

Un poco de historia

Primeros repertorios poéticos

POSTDATA

- Gaston Raynaud, *Bibliographie des chansonniers français des XIIIe et XIVe siècles,* 1884.
- Gotthold Naetebus, *Die nicht lyrischen Strophenformen des Altfranzösischen*, 1891.
- Pillet & Carstens, *Bibliographie des Troubadours,* 1933

VIII. aaaa. 12 S.

1. La vie et l'istoire du mauvais riche ho

Devant l'uis au riche home le ladre s'arresta
Por la grant fain qu'il ot forment se dementa.

Hs. Paris, Nat. bibl. f. fr. 957 Bl. 118.

Mit den ersten 5 Zeilen angeführt 1848 von P. Paris, Mss. fr. VII, 339. — Über ein in paarweis reimenden Achtsilbnern verfafstes Livre de l'exemple du riche homme et du ladre vgl. van Hamel, Rencl. S. CXCIX und Rom. XVIII, 642.

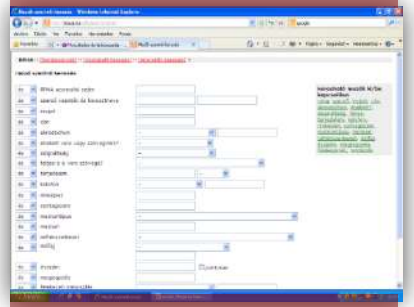
Hs.: An

Concordance avec les répertoires:	**Långfors : 90** **Raynaud - Spanke : -** **Mölk - Wolfzettel : -**
Incipit:	**Devant l'uys au riche homme le ladre trespassa**
Variantes de l'incipit:	Devant l'uys au riche homme le ladre trespassa, Devant l'uys au riche homme le ladre s'aresta,
Source:	Cambridge, Magdalene College, Coll. S. Pepys 1938 Paris, BNF fr. 957
Titre:	
Auteur:	(texte anonyme) Cambridge, Magdalene College, Coll. S. Pepys 1938
Auteurs secondaires:	
Versification:	**texte entièrement en vers** **poème strophique** **poème isostrophique**
Nombre des strophes:	28 Cambridge, Magdalene College, Coll. S. Pepys 1938 Paris, BNF fr. 957
Structure strophique:	12aaaa
Irrégularités métriques (numéro de la strophe – structure strophique):	8 - 12aa Paris, BNF fr. 957 Cambridge, Magdalene College, Coll. S. Pepys 1938 10 - 12aaa Paris, BNF fr. 957 Cambridge, Magdalene College, Coll. S. Pepys 1938 16 - 12aaa Paris, BNF fr. 957 Cambridge, Magdalene College, Coll. S. Pepys 1938

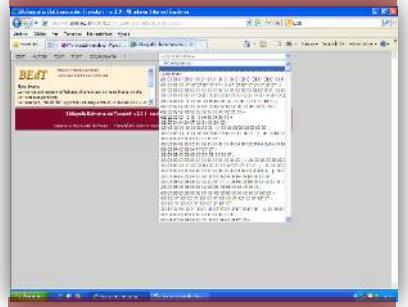
El puzzle de repertorios dig

POSTDATA
Poetry Standardization and Linked Open Data

RPHA - Hungary

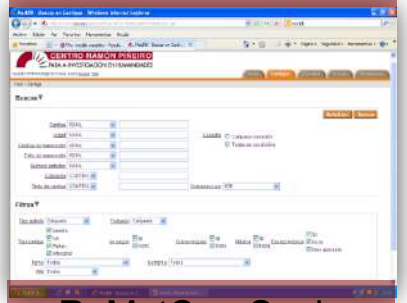


BedT - Italy

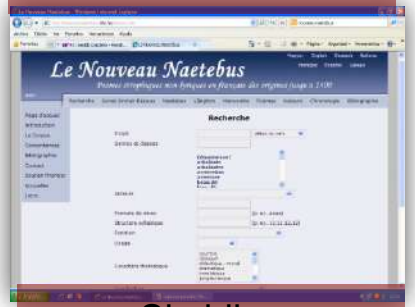


Cantigas Oxford

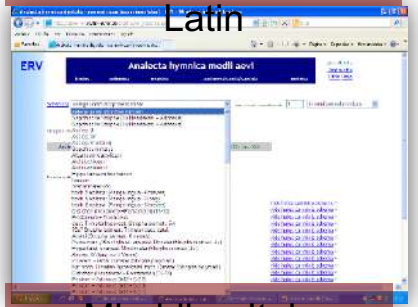
MedDB2 – Galicia



Naetebus - France



Analecta Hymnica Latin



ReMetCa - Spain

Skaldic - Scandinavia

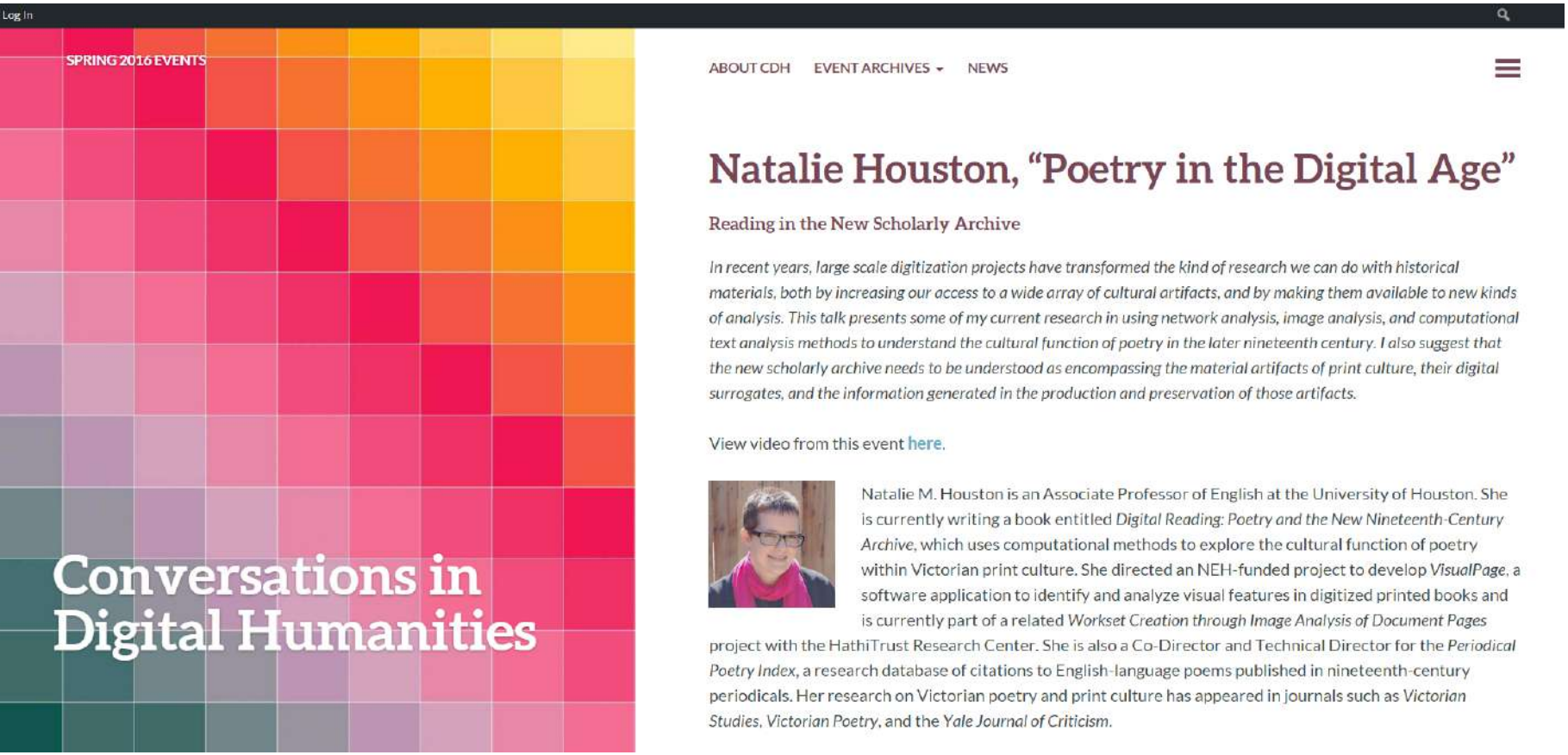
Mirabile - Italy

Barreras tecnológicas

POSTDATA
Poetry Standardization and Linked Open Data

¿Cómo organizar los datos?

POSTDATA
Poetry Standardization and Linked Open Data



- http://blogs.cornell.edu/dhconversations/natalie-

Visión de POSTDATA: facilitar la investigación de los humanistas mediante la transformación digital

POSTDATA
Poetry Standardization and Linked Open Data

Investigación tradicional		Humanidades Digitales
Contar sílabas con los dedos	→	Contar sílabas automáticamente
Comparar 2-3 poemas al mismo tiempo	→	Comparar miles de poemas a la vez
Buscar testimonios en libros y artículos	→	Buscar múltiples evidencias a través de un motor de búsqueda
Investigación individual	→	Investigación colaborativa

Las barreras filológicas se deben a distintas formas de conceptualizar

POSTDATA
Poetry Standardization and Linked Open Data

Estuans intrinsecus ira vehementi
in amaritudine loquor mee menti.
factus de materia levis elementi
4 folio sum similis, de quo ludunt venti.

Cum sit enim proprium viro sapienti,
supra petram ponere sedem fondamenti,
stultus ego comparor fluvio labenti,
8 sub eodem aere numquam permanenti.

Feror ego veluti sine nauta navis,
ut per vias aeris vaga fertur avis;
non me tenent vincula, non me tenet clavis,
12 quero mei similes et adiungor pravis.

Carmina Burana, 191

Alejandrino

Goliárdico

12A12A12A12A
(Tradición romance)

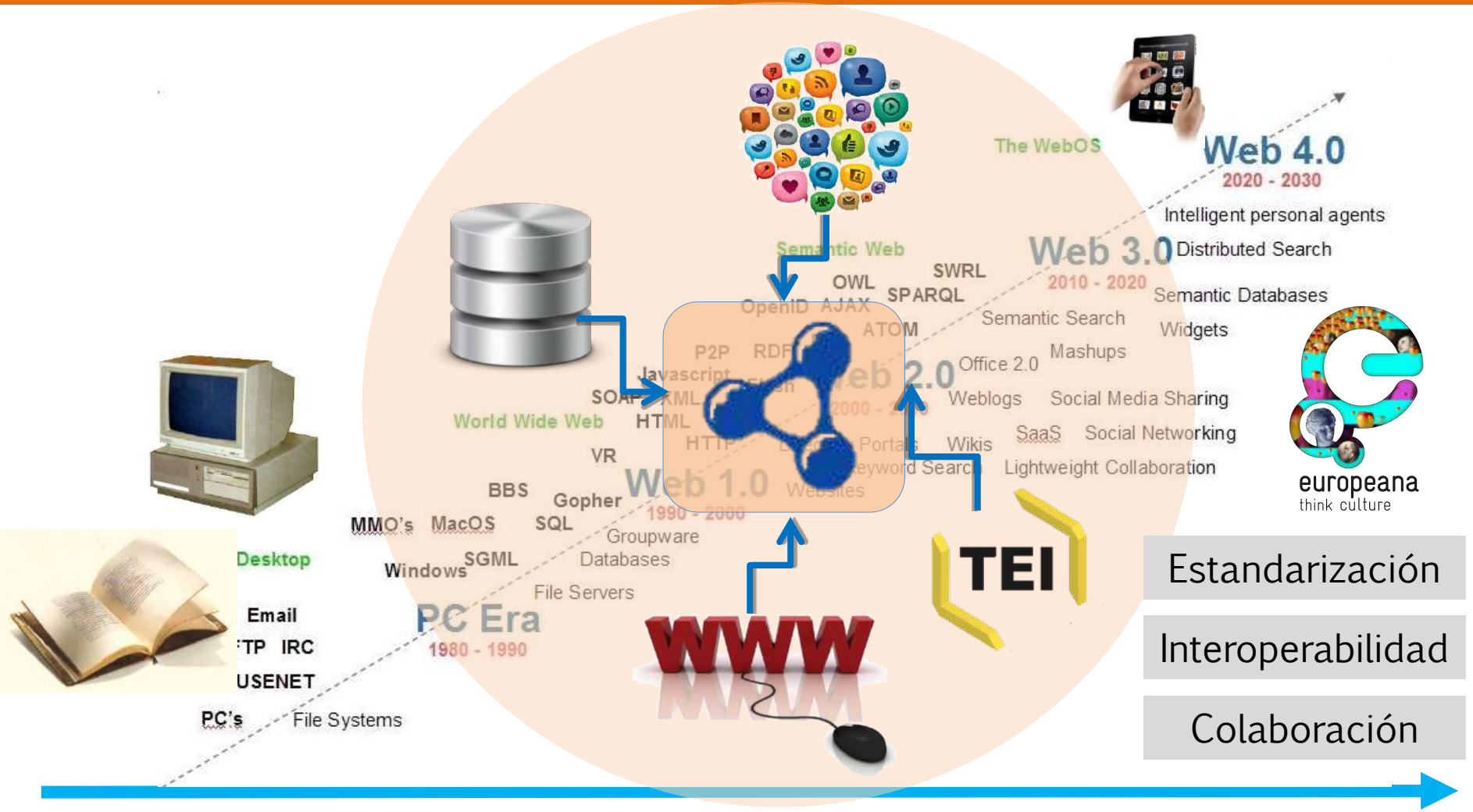
4x(7pp+7p)
(Latín Clásico)

Las barreras tecnológicas por falta de estandarización e interoperabilidad

POSTDATA
Poetry Standardization and Linked Open Data

POSTDATA es la clave para enlazar tecnologías haciéndolas interoperables

POSTDATA
Poetry Standardization and Linked Open Data



POSTDATA: un modelo abstracto con elementos comunes entre los diferentes sistemas poéticos

POSTDATA
Poetry Standardization and Linked Open Data

DBpedia

GeoNames

MusicBrainz

Author
Bibliographical_references
Manuscript
Poem
Line
Versification_system
Thing
Edition
Stanza
Syllable
URL
Date
Complete_work

Latin poetry

Spanish poetry

Galician poetry

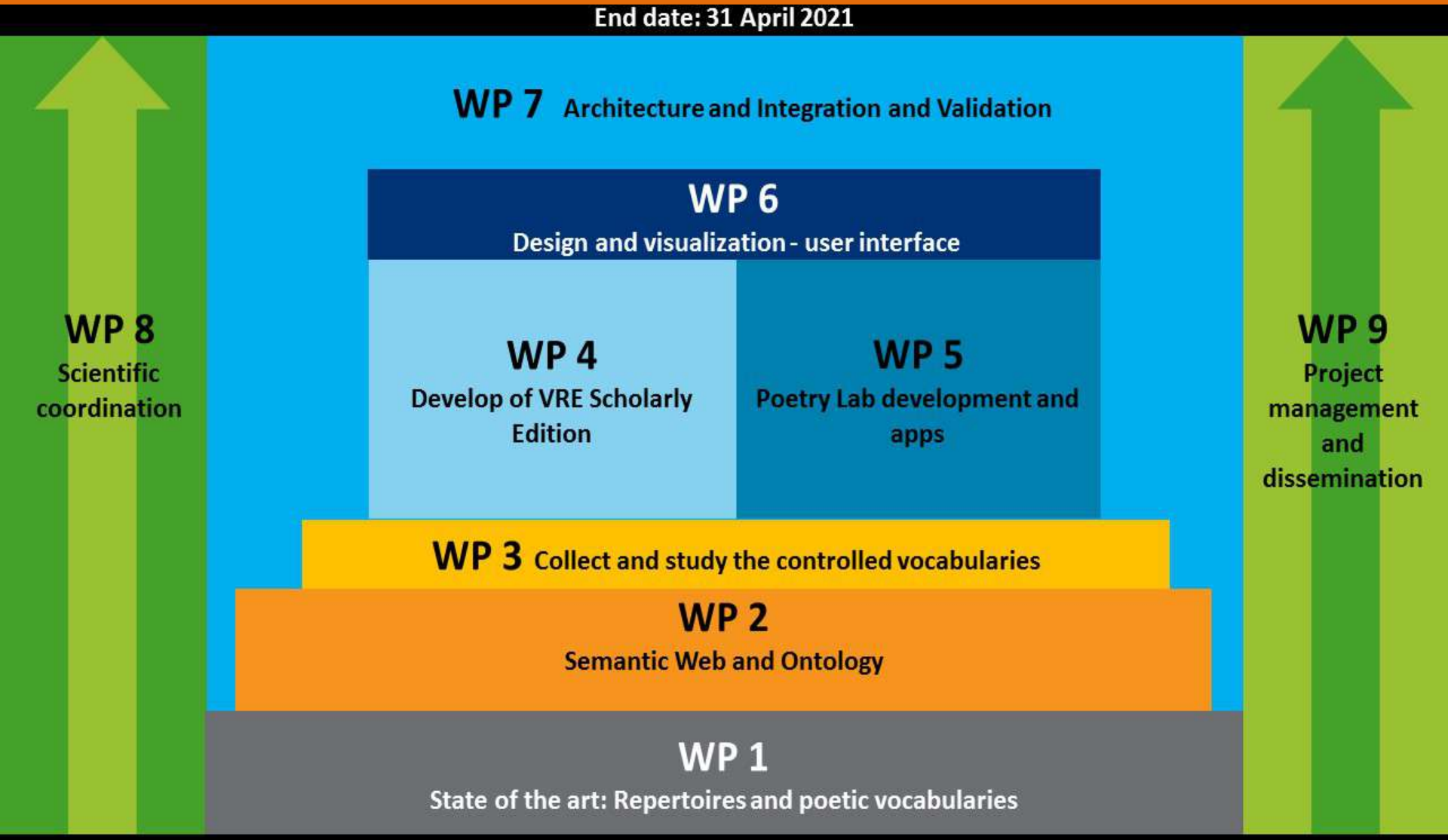
Italian poetry

French poetry

Occitan poetry

Los Workpackages de POSTDATA

POSTDATA
Poetry Standardization and Linked Open Data



Los tres ejes de POSTDATA

POSTDATA
Poetry Standardization and Linked Open Data

Infraestructura semántica

Datos Abiertos Enlazados (LOD)

Entorno virtual de investigación

Edición digital académica

Poetry Lab

Procesamiento del lenguaje natural NLP

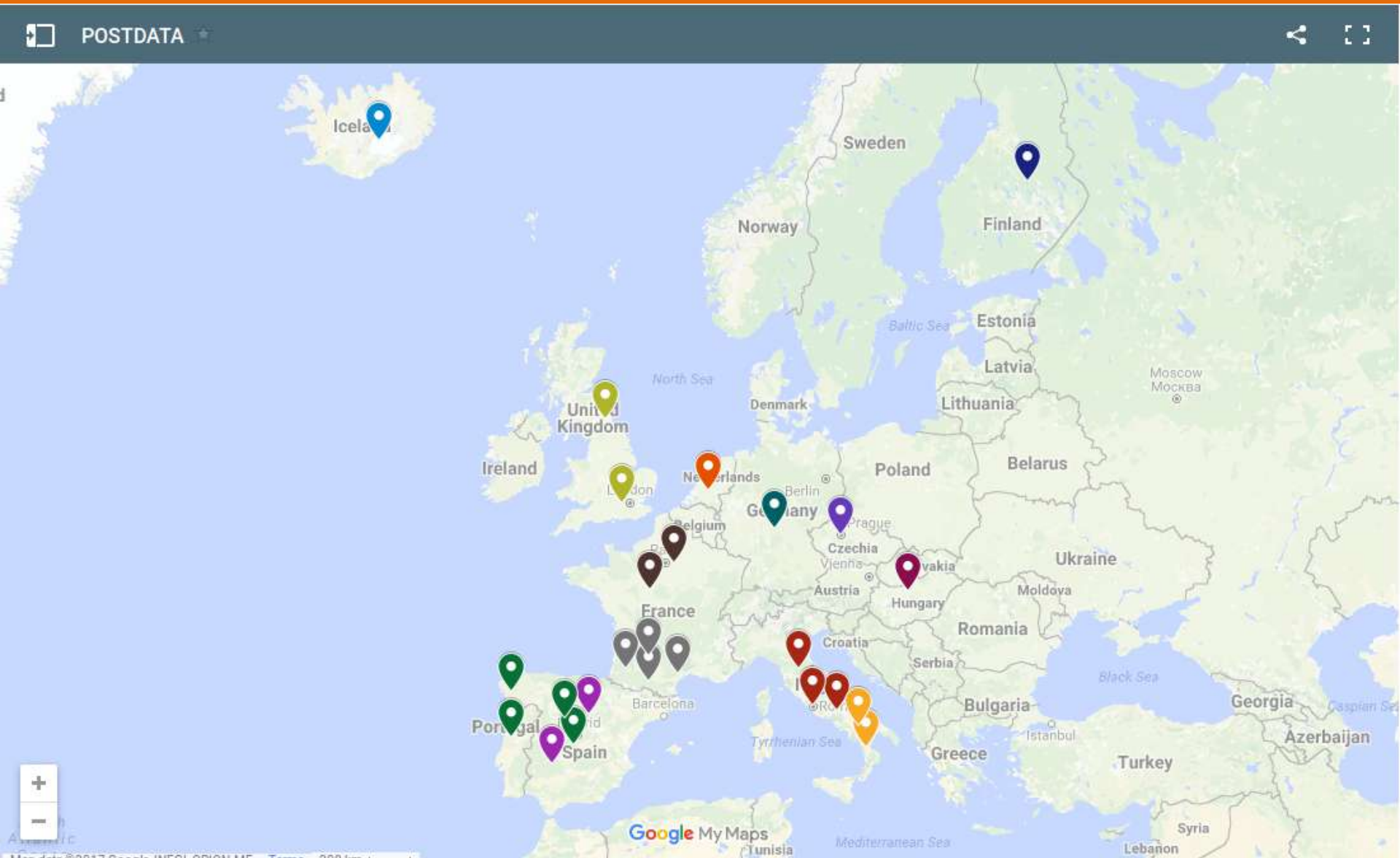
Dónde estamos... las bases de datos

POSTDATA
Poetry Standardization and Linked Open Data

- 28 bases de datos contactadas
- 26 respuestas positivas de colaboración con POSTDATA
- 11 databases se han utilizado para definir el Domain Model V0.1
- Seguimos abiertos a encontrar nuevos repertorios...
- http://postdata.linhd.es/partners/

The databases - partners

POSTDATA
Poetry Standardization and Linked Open Data

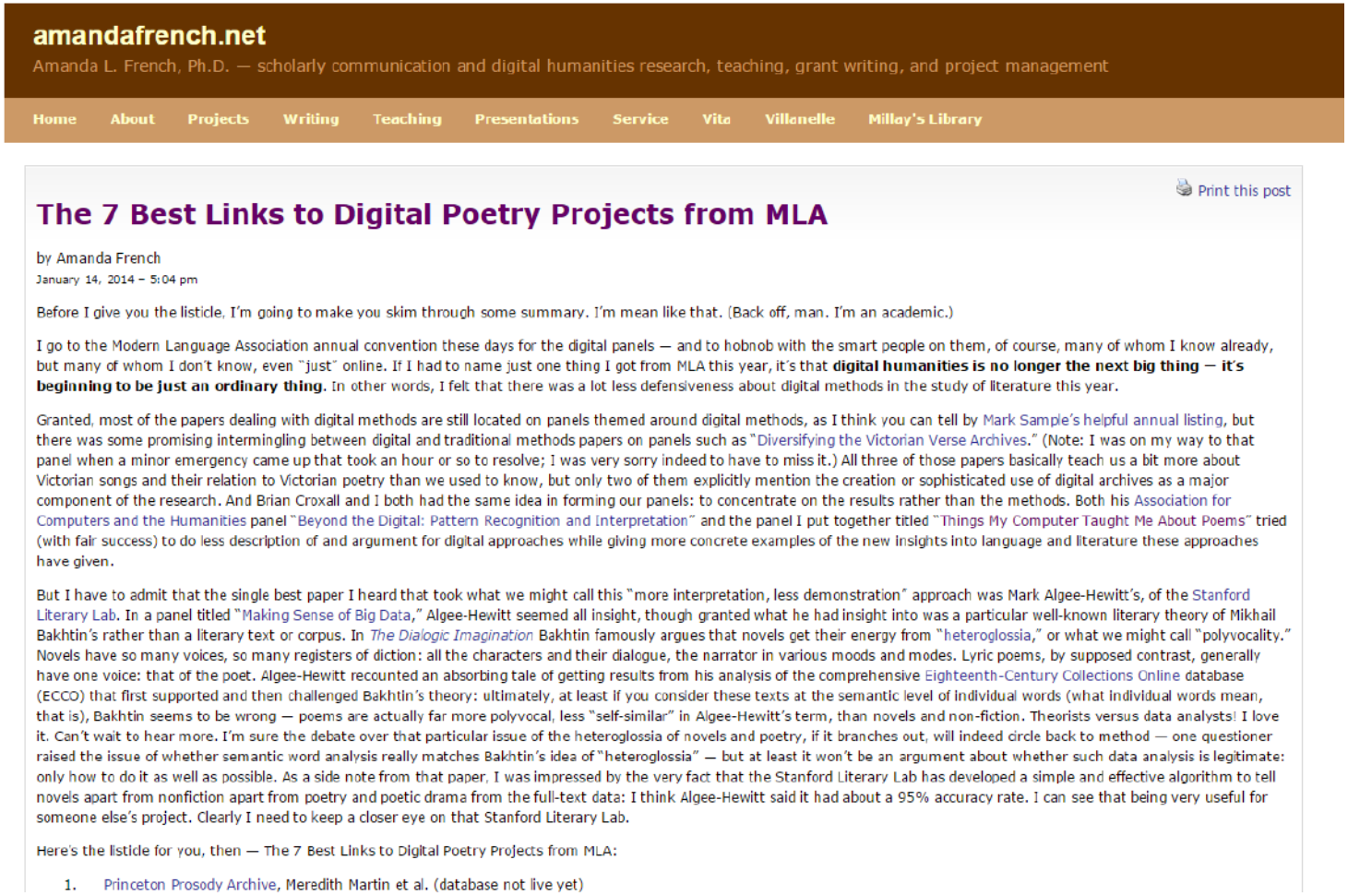


elena@coverwallet.com
http://postdata.linhd.es
@linhduned

POSTDATA

Poetry Standardization
and Linked Open Data

How do we organize data?



amandafrench.net

Amanda L. French, Ph.D. — scholarly communication and digital humanities research, teaching, grant writing, and project management

Home About Projects Writing Teaching Presentations Service Vita Villanelle Millay's Library

Print this post

The 7 Best Links to Digital Poetry Projects from MLA

by Amanda French

January 14, 2014 – 5:04 pm

Before I give you the listicle, I'm going to make you skim through some summary. I'm mean like that. (Back off, man. I'm an academic.)

I go to the Modern Language Association annual convention these days for the digital panels — and to hobnob with the smart people on them, of course, many of whom I know already, but many of whom I don't know, even "just" online. If I had to name just one thing I got from MLA this year, it's that **digital humanities is no longer the next big thing — it's beginning to be just an ordinary thing.** In other words, I felt that there was a lot less defensiveness about digital methods in the study of literature this year.

Granted, most of the papers dealing with digital methods are still located on panels themed around digital methods, as I think you can tell by Mark Sample's helpful annual listing, but there was some promising intermingling between digital and traditional methods papers on panels such as "Diversifying the Victorian Verse Archives." (Note: I was on my way to that panel when a minor emergency came up that took an hour or so to resolve; I was very sorry indeed to have to miss it.) All three of those papers basically teach us a bit more about Victorian songs and their relation to Victorian poetry than we used to know, but only two of them explicitly mention the creation or sophisticated use of digital archives as a major component of the research. And Brian Croxall and I both had the same idea in forming our panels: to concentrate on the results rather than the methods. Both his Association for Computers and the Humanities panel "Beyond the Digital: Pattern Recognition and Interpretation" and the panel I put together titled "Things My Computer Taught Me About Poems" tried (with fair success) to do less description of and argument for digital approaches while giving more concrete examples of the new insights into language and literature these approaches have given.

But I have to admit that the single best paper I heard that took what we might call this "more interpretation, less demonstration" approach was Mark Algee-Hewitt's, of the Stanford Literary Lab. In a panel titled "Making Sense of Big Data," Algee-Hewitt seemed all insight, though granted what he had insight into was a particular well-known literary theory of Mikhail Bakhtin's rather than a literary text or corpus. In *The Dialogic Imagination* Bakhtin famously argues that novels get their energy from "heteroglossia," or what we might call "polyvocality." Novels have so many voices, so many registers of diction: all the characters and their dialogue, the narrator in various moods and modes. Lyric poems, by supposed contrast, generally have one voice: that of the poet. Algee-Hewitt recounted an absorbing tale of getting results from his analysis of the comprehensive Eighteenth-Century Collections Online database (ECCO) that first supported and then challenged Bakhtin's theory: ultimately, at least if you consider these texts at the semantic level of individual words (what individual words mean, that is), Bakhtin seems to be wrong — poems are actually far more polyvocal, less "self-similar" in Algee-Hewitt's term, than novels and non-fiction. Theorists versus data analysts! I love it. Can't wait to hear more. I'm sure the debate over that particular issue of the heteroglossia of novels and poetry, if it branches out, will indeed circle back to method — one questioner raised the issue of whether semantic word analysis really matches Bakhtin's idea of "heteroglossia" — but at least it won't be an argument about whether such data analysis is legitimate: only how to do it as well as possible. As a side note from that paper, I was impressed by the very fact that the Stanford Literary Lab has developed a simple and effective algorithm to tell novels apart from nonfiction apart from poetry and poetic drama from the full-text data: I think Algee-Hewitt said it had about a 95% accuracy rate. I can see that being very useful for someone else's project. Clearly I need to keep a closer eye on that Stanford Literary Lab.

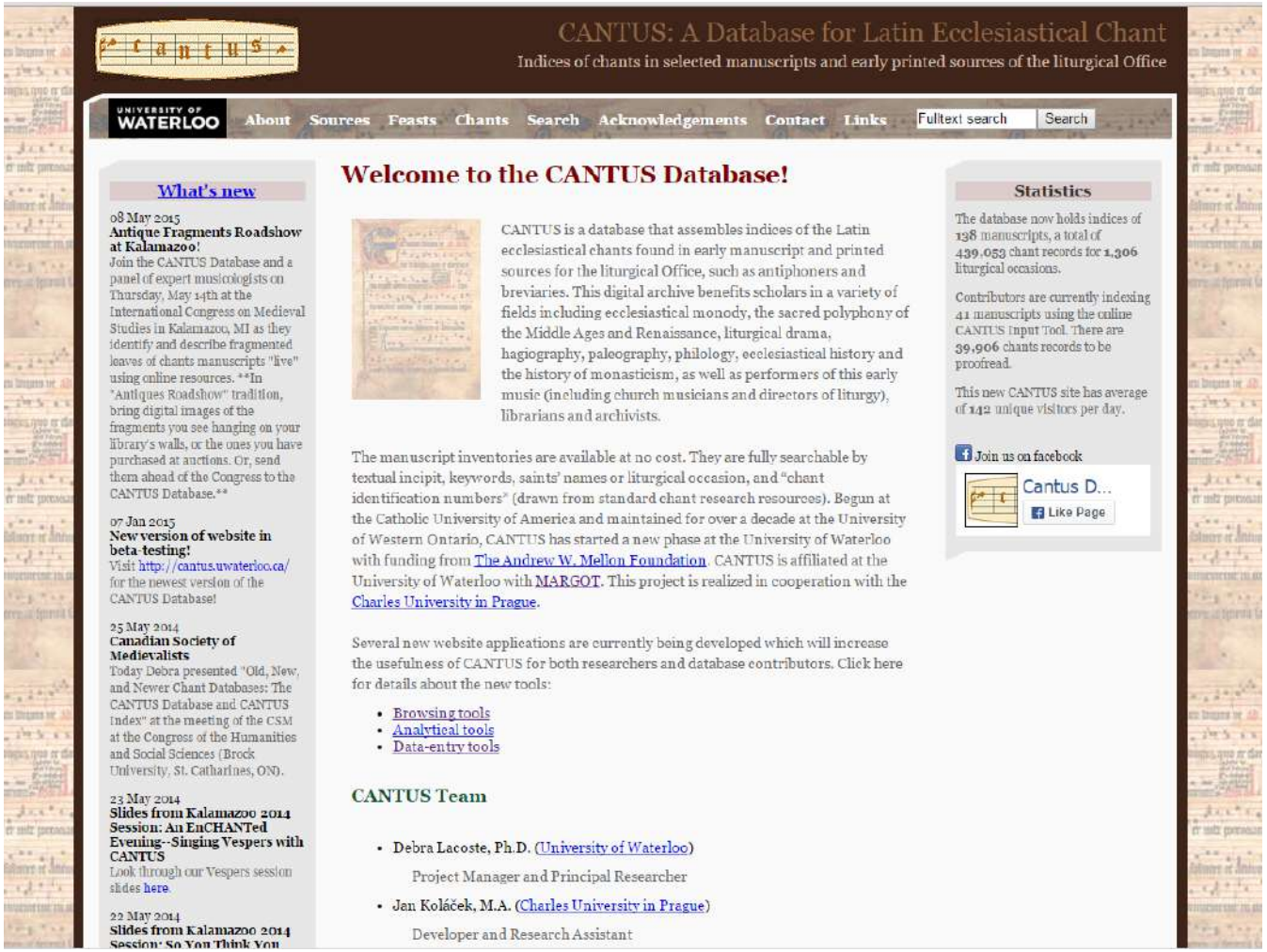
Here's the listicle for you, then — The 7 Best Links to Digital Poetry Projects from MLA:

1. Princeton Prosody Archive, Meredith Martin et al. (database not live yet)

A typology of poetry projects

1. Repositories of poetic corpora and collections
   - Images
   - Texts
   - Music
2. Digital editions + XML markup
3. Quantitative analysis: Statystical analysis, NLP and data mining
4. Experimental projects: Computational creativity
   - Other disciplines: aesthetics, psychology
5. Visualization experiments
6. Social + user-oriented projects
7. Semantic web-based projects

1. Repositories and databases

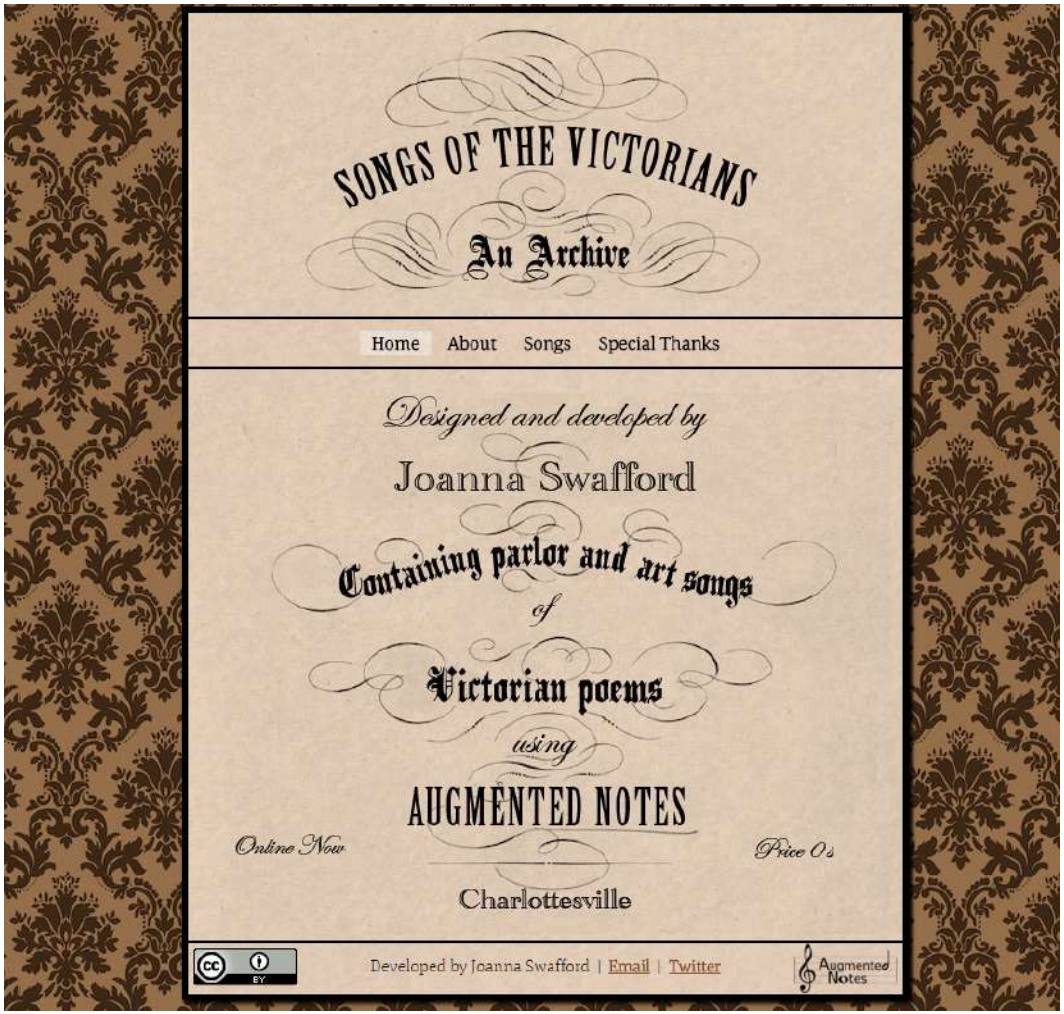


http://cantusdatabase.org/

1. Repositories and databases

http://vm136.lib.berkeley.edu/BANC/philobiblon/searchwork_en.html

1. Repositories and databases



http://www.songsofthevictorians.com/#home

1. Repositories and databases

Search

American Poetry of the Great War

A Digital Archive

Home | About the Archive | Volumes | Authors | Blog | Bibliography

Welcome

The American Poetry of the Great War Digital Archive is intended as a central respository of poetry written by Americans during and immediately following World War I.

We hope the archive will be a resource for those interested in exploring the war, along with Americans' responses to it, through literature.

News

- New Additions to the Archive
- The Original Vision

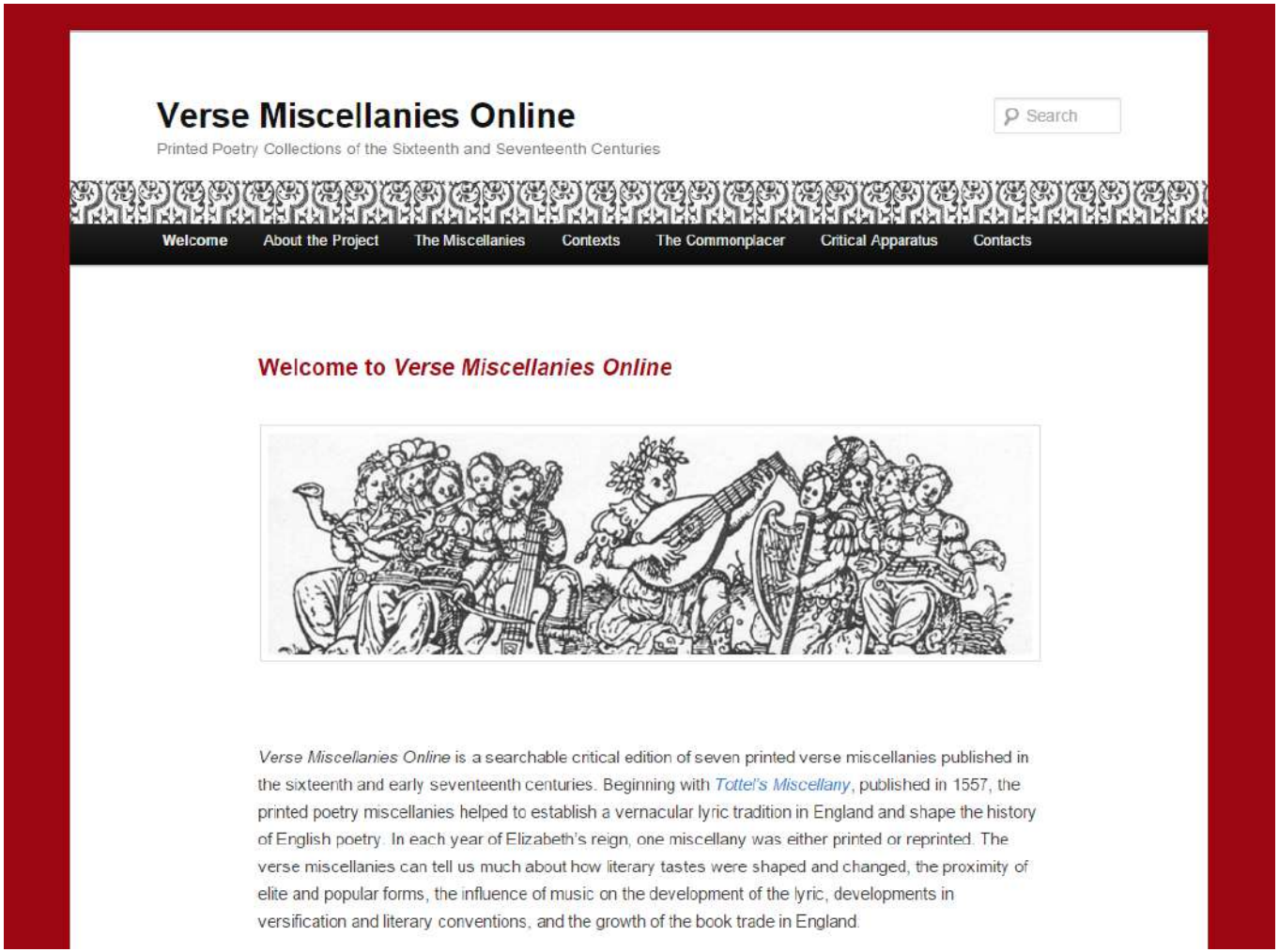
More

Embarked for France. Western Newspaper Union, 1917

http://americanww1poetry.drupalgardens.com/

2. Digital editions using XML-TEI

2. Digital editions using XML-TEI



http://versemiscellaniesonline.bodleian.ox.ac.uk/

3.Quantitative analysis: Statystical methods, NLP and data mining

<oo>→<met> Meter, rhythm, and rhyme

Maintained by: David J. Birnbaum (djbpitt@gmail.com)
Last modified: 2015-05-30T19:17:57+0000

The computationally assisted analysis of formal features in Russian poetry

About this site

This project, developed by Elise Thorsen and David J. Birnbaum, is dedicated to the design and deployment of computational tools to aid the analysis of formal features in Russian poetry. These tools involve the detection of stress in Russian texts presented in native orthography and the subsequent derivation of meter and rhyme patterns from stress and orthographic information. The output of those processes is an XML document that can provide the basis for visualizations of formal features (see, for example, the verse tables of the sample poems below). Beyond these reports, the informational mark-up of poetic texts is intended to enable the characterization of formal features in large digitized poetic corpora to support research in sub-fields including quantitative metrics, genre studies, and diachronic studies of stylistics and influence.

Research reports

- Enabling the automated identification and analysis of meter and rhyme in Russian verse
  - Proposal, DH2015: Global digital humanities (Sydney)
  - Slides from presentation at the LAUDATIO mini-workshop, Berlin 2013-05-07
- Perspectives on computer programming for the humanities. By Lawrence D. Adams and David J. Birnbaum; originally published in *Text technology* 7, 1 (1997), 1–17, and updated for the present online version.
- Strong and weak metrical positions

Sample poems

Under development; interim results

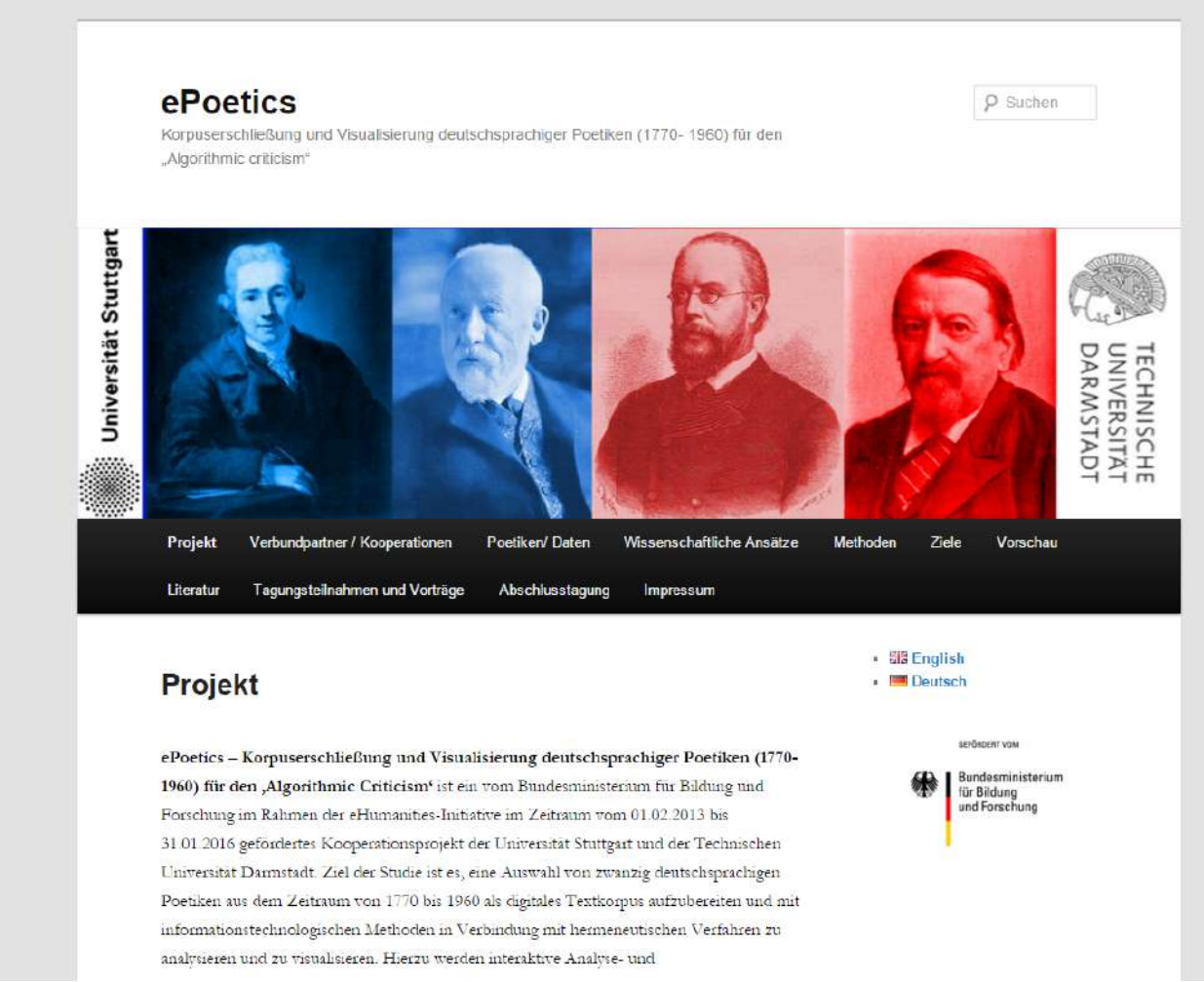
- Ахматова, Анна Андреевна: Вечером, Лилии, Любовь, Наследница, Он любил
- Гиппиус, Зинаида Николаевна: Нелюбовь
- Пастернак, Борис Леонидович: Его речь
- Пушкин, Александр Сергеевич: Бесы, Евгений Онегин, К *** (Ты богоматерь, нет сомненья), К *** (Я помню чудное мгновенье), Кавказский пленник, Пророк

Miscellaneous

- Bibliography
- Internet resources
- A computational approach to identifying the place of stress in Russian poetry [watch this space]
- Early mockup output

http://poetry.obdurodon.org/

3.Quantitative analysis: Statystical methods, NLP and data mining

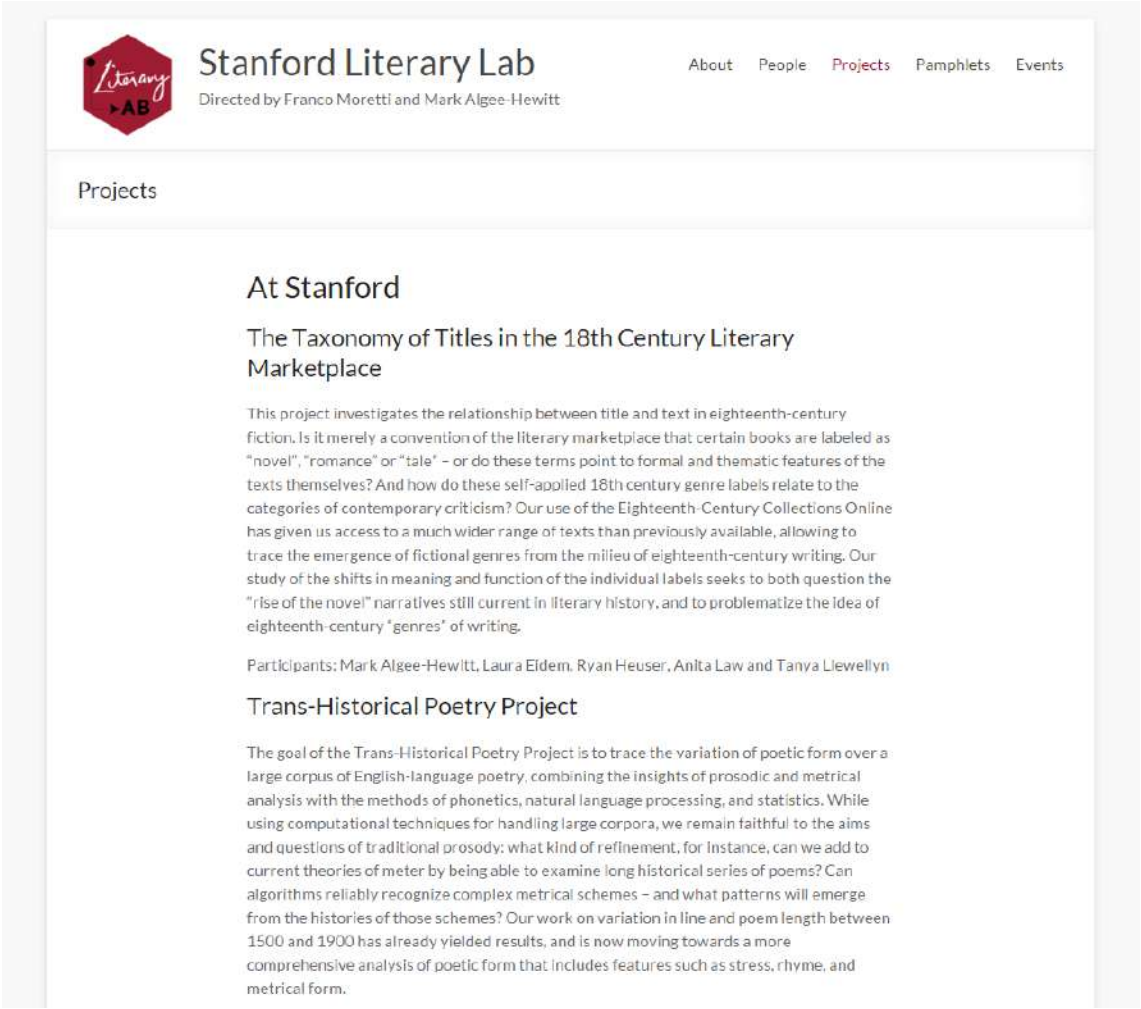


www.epoetics.de

3.Quantitative analysis: Statystical methods, NLP and data mining

http://www.crisco.unicaen.fr/verlaine/presentation.html

3.Quantitative analysis: Statystical methods, NLP and data mining



Stanford Literary Lab

Directed by Franco Moretti and Mark Algee-Hewitt

About People Projects Pamphlets Events

Projects

At Stanford

The Taxonomy of Titles in the 18th Century Literary Marketplace

This project investigates the relationship between title and text in eighteenth-century fiction. Is it merely a convention of the literary marketplace that certain books are labeled as "novel", "romance" or "tale" – or do these terms point to formal and thematic features of the texts themselves? And how do these self-applied 18th century genre labels relate to the categories of contemporary criticism? Our use of the Eighteenth-Century Collections Online has given us access to a much wider range of texts than previously available, allowing to trace the emergence of fictional genres from the milieu of eighteenth-century writing. Our study of the shifts in meaning and function of the individual labels seeks to both question the "rise of the novel" narratives still current in literary history, and to problematize the idea of eighteenth-century "genres" of writing.

Participants: Mark Algee-Hewitt, Laura Eidem, Ryan Heuser, Anita Law and Tanya Llewellyn

Trans-Historical Poetry Project

The goal of the Trans-Historical Poetry Project is to trace the variation of poetic form over a large corpus of English-language poetry, combining the insights of prosodic and metrical analysis with the methods of phonetics, natural language processing, and statistics. While using computational techniques for handling large corpora, we remain faithful to the aims and questions of traditional prosody: what kind of refinement, for instance, can we add to current theories of meter by being able to examine long historical series of poems? Can algorithms reliably recognize complex metrical schemes – and what patterns will emerge from the histories of those schemes? Our work on variation in line and poem length between 1500 and 1900 has already yielded results, and is now moving towards a more comprehensive analysis of poetic form that includes features such as stress, rhyme, and metrical form.

https://litlab.stanford.edu/current-projects/

4. Computational creativity

Natural Interaction based on Language

Search this site: Search

Navigation
- Main
- Members
- Research
- Projects
- Publications
- Contact
- Login

Home

WASP

WASP (the Wishful Automatic Spanish Poet), is a rule-based system that generates automatically verses in Spanish. It is actually, a family of programs that generate formal Spanish Poetry. A pedagogical application that is our current benchmark for exploring Computational Creativity in machine use of language. Builds correct formal Spanish poetry starting from very basic metric information. As more and more language resources are developed, additional levels of linguistic information slowly creep in (morphology, syntax, semantics...).

[31st May 2005] The WASP family of automatic poets is presented at the Feria del Libro de Madrid. A brief presentation on poetry generated by computers was followed by a live demonstration of how the WASPos convert a message provided by the user into formal Spanish verse. An informal chat followed, with the participation of (human) poet María Sáinz.

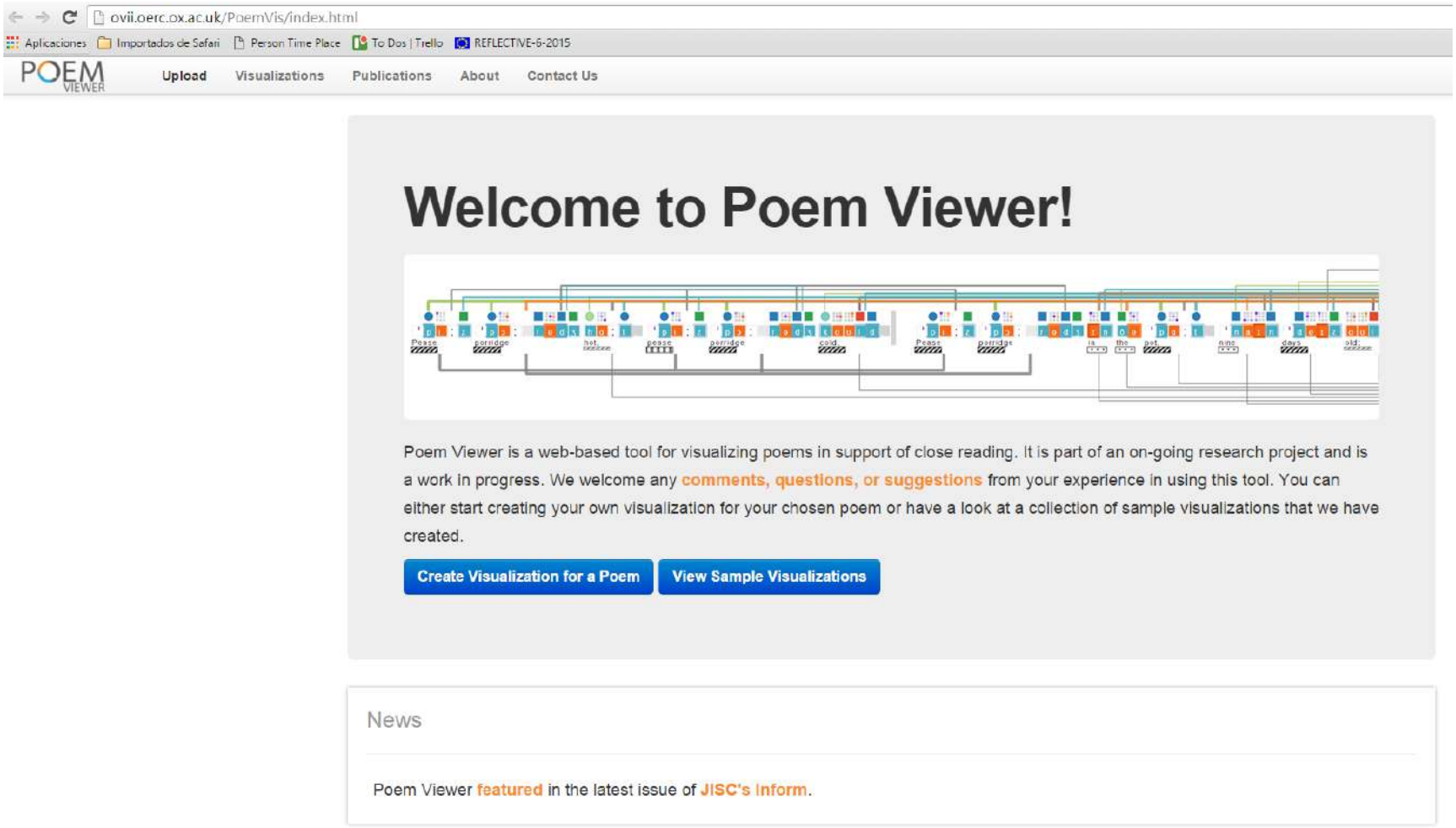
This is what the press had to say about the event:

- El Pais digital, 8th June 2005
- Libertad Digital, 10th June 2005
- CRTVG, 15th April 2010

You can find more information about WASP related publications in Publications.

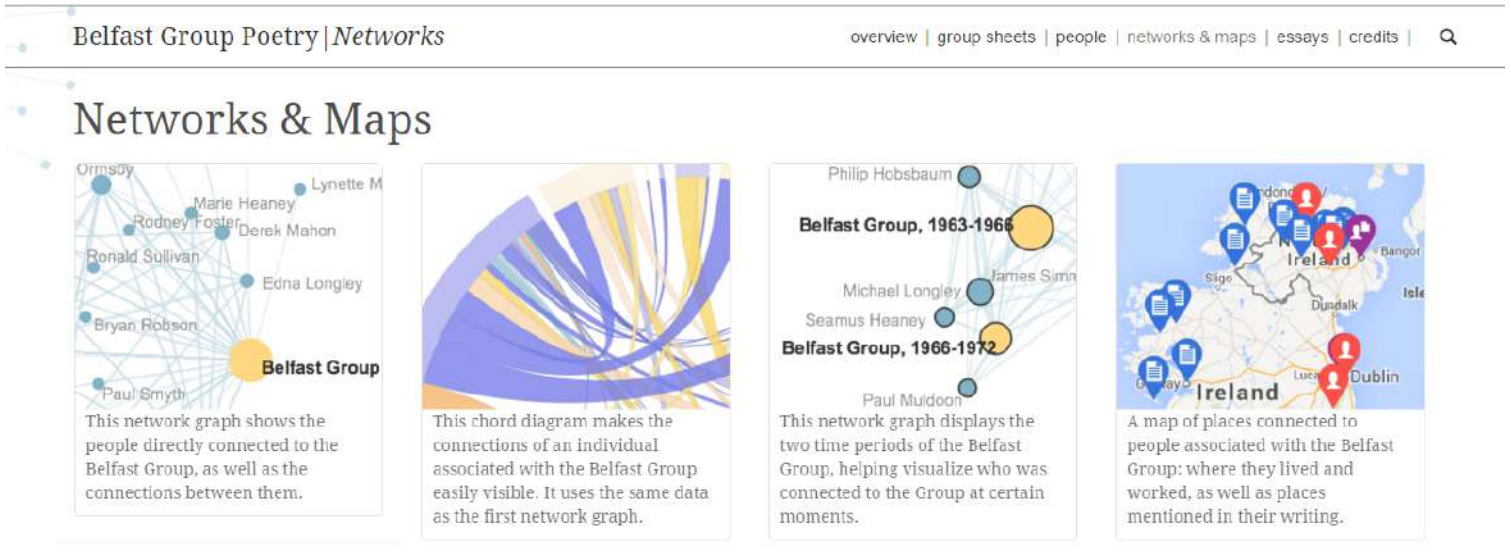
http://nil.fdi.ucm.es/index.php?q=node/206

5.Visualization-oriented projects



http://ovii.oerc.ox.ac.uk/PoemVis/index.html

5.Visualization-oriented projects



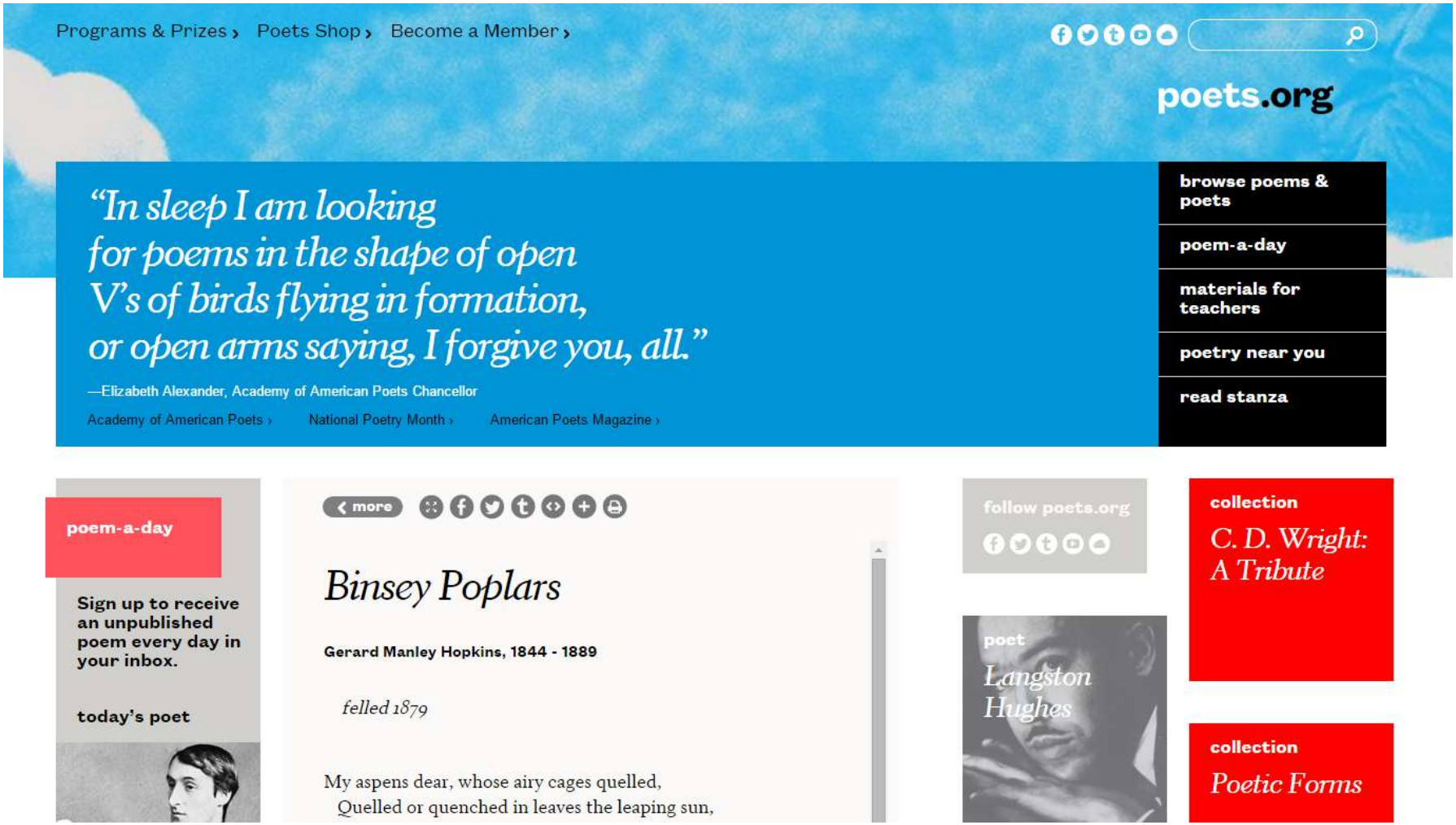
After the Belfast Group stopped meeting in 1972, many participants downplayed its importance to the development of their writing. Regardless of its ultimate impact, the writing workshop nevertheless connected many authors in Northern Ireland. The pages linked above visualize these networks of relationships, as well as the writers' connections to specific places.

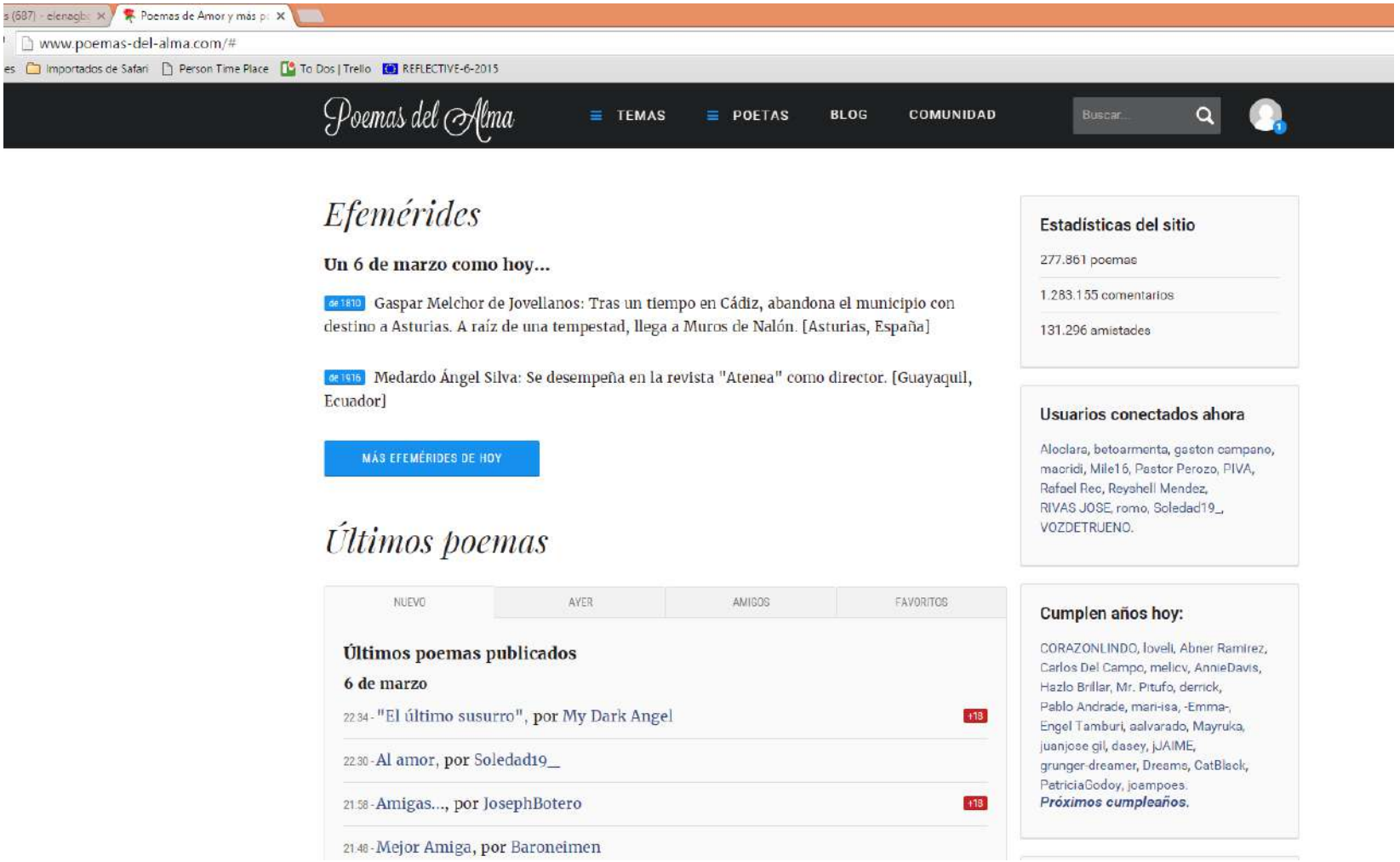
The data used to create these visualizations come from the digitized drafts of Group sheets as well as the poets' collected materials in Emory University's Manuscript, Archives, and Rare Book Library (MARBL). The Namedropper software developed for Belfast Group Poetry | *Networks* helped the team tag people and places mentioned in the archival descriptions or poetry with unique identifiers. The computer can now "understand" that Seamus Heaney wrote to Michael Longley and Derek Mahon, as well as writing frequently about County Tyrone. As a result, this "understanding" can be quantified and represented graphically. (A more technical description of the RDF-generation process is available.) The visualizations were created to represent different views on the relationships among Belfast Group participants, including how connected they were to the writing workshop itself. You can also see the network of connections among individual writers and their peers on individual biography pages.

This approach allows the investigation of the Belfast literary circle in a way that would be very difficult through regular archival research, especially since the project draws on catalogue information about letters that are currently sealed to researchers. Nevertheless, the visualizations cannot provide a complete picture. Not all of the poetry Group sheets have been digitized nor do MARBL's collections hold every letter sent by or to a poet, let alone materials from every author in the Belfast Group. That said, even incomplete visualizations provide a new way to investigate the formation of a literary community.

http://belfastgroup.digitalscholarship.emory.edu/network/

6.Social & user-oriented projects



www.poets.org

6.Social & user-oriented projects