

PLN, ciencia auxiliar para las humanidades digitales

Gimena del Rio Riande, (CONICET), Argentina

Nidia Hernández, (CONICET), Argentina



Humanidades Digitales

- u Campo, disciplina, método, metodología.
- u Trabajo interseccional (Filología, Literatura, Historia, Lingüística, Bibliotecología, Ciencias de la Información, Programación).
- u Trabajo colaborativo.
- u Concepción extendida de los términos “humanidades” y “digital”.
- u Utilizan materiales digitalizados o born-digital combinando métodos/metodologías de las humanidades y ciencias sociales con herramientas y métodos informáticos (codificación, visualización de datos, recuperación de información, minería de datos, estadística, minería de textos) y publicación digital.

Generar nuevo conocimiento humanístico
utilizando herramientas digitales.





Machine Translation
IBM



Index Thomisticus.
Roberto Busa



*Computerized
Concordance to the
Revised Standard
Version of the Bible*
John W. Ellison



CORPUS THOMISTICUM
INDEX THOMISTICUS
by Roberto Busa SJ and associates
web edition by Eduardo Bernot and Enrique Alarcón
English version

Search:

concordances

terms

works

options

new search

FOUND 106 CASES IN 62 PLACES

1-10



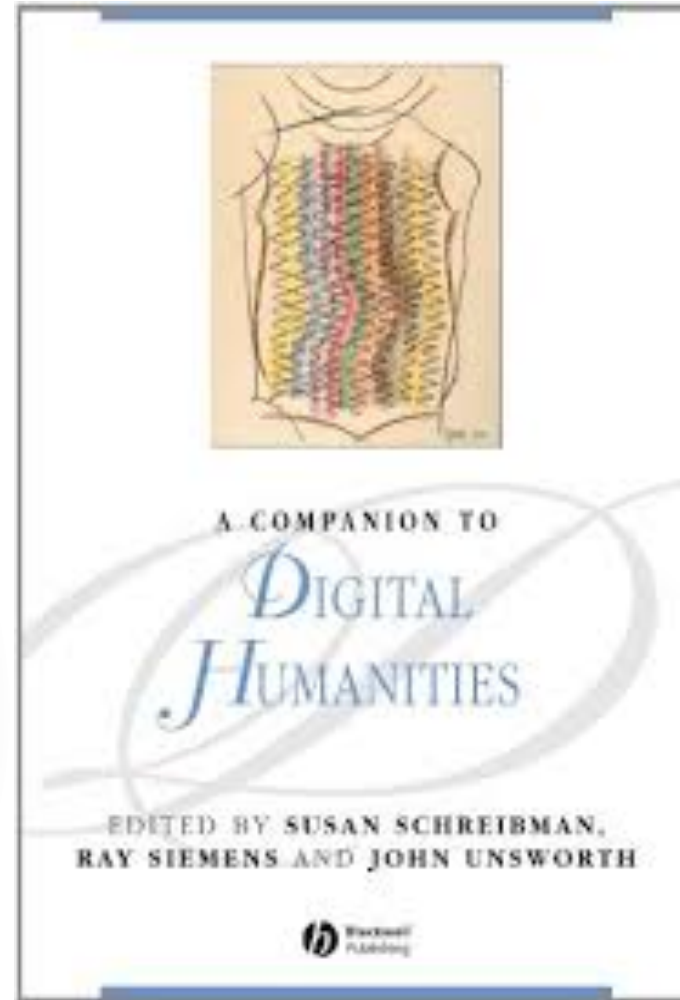
CASE 1. PLACE 1. *Super Sent., lib. 3 d. 21 q. 2 a. 3 ad 2.* Ad secundum dicendum, quod argumenta illa sufficienter probabant illud ad quod probandum inducebantur, ut dictum est. Sed fides non est directe de illo quod probabatur sicut de objecto, sed de divinitate, quae probari non potest homini in via: unde **Thomas**, de quo dictum est, Joan. 20, 29, *quia vidisti, credidisti*: vidit hominem, et Deum confessus est; et propter hoc argumenta illa ex suppositione fidei procedebant, quod scilicet homo ille, in quo naturae veritas, vel aliquid hujusmodi ostendebatur, esset Deus; unde non tollebat neque meritum neque ratio fidei.

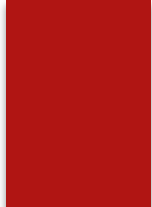
CASE 2. PLACE 2. *Super Sent., lib. 3 d. 21 q. 2 a. 4 qc. 3 arg. 3.* Praeterea, Hieronymus dicit ad Marcellam, quod non meretur tangere Christum qui resurrectionem ejus non credit; propter quod dictum est Magdalenae, Joan. 20, 17: *noli me tangere*. Sed **Thomas** resurrectionem non credebat. Ergo Christus cicatrices vulnerum ei palpabiles exhibere non debuit.

CASE 3. PLACE 3. *Super Sent., lib. 3 d. 21 q. 2 a. 4 qc. 3 ad 3.* Ad tertium dicendum, quod **Thomas** ad solum aspectum vulnerum credit; unde dictum est ei, Joan. 20, 29: *quia vidisti me, credidisti*. Sed tamen tetigit eum, ut a nobis omnem dubitationem excluderet: quia testis erat futurae resurrectionis; unde Leo Papa dicit: *suffecerat ad fidem propriam quod viderat; sed nobis operatus est, ut tangeret quod videbat*. [...] ¹

<http://www.corpusthomisticum.org/it>

A Companion to Digital Humanities





- 16 **7. Linguistics Meets Exact Sciences**
- 8. [Literary Studies](#)
- 9. [Music](#)
- 10. [Multimedia](#)
- 11. [Performing Arts](#)
- 12. ["Revolution? What Revolution?" Successes and Limits of C](#)

- 24 **Part II: Principles**
- 1 **13. [How the Computer Works](#)**
- 7 **14. [Classification and its Structures](#)**
- 15. [Databases](#)
- 13 **16. [Marking Texts of Many Dimensions](#)**
- 2 **17. [Text Encoding](#)**
- 1 **18. [Electronic Texts: Audiences and Purposes](#)**
- 19. [Modeling: A Study in Words and Meanings](#)

- 16 **Part III: Applications**
- 20. [Stylistic Analysis and Authorship Studies](#)
- 7 **21. [Preparation and Analysis of Linguistic Corpora](#)**
- 1 **22. [Electronic Scholarly Editing](#)**
- 23. [Textual Analysis](#)
- 24. [Thematic Research Collections](#)
- 25. [Print Scholarship and Digital Resources](#)
- 26. [Digital Media and the Analysis of Film](#)
- 2 **27. [Cognitive Stylistics and the Literary Imagination](#)**
- 1 **28. [Multivariant Narratives](#)**
- 29. [Speculative Computing: Aesthetic Provocations in Human](#)
- 5 **30. [Robotic Poetics](#)**

- 1 **Part IV: Production, Dissemination, Archiving**
- 31. [Designing Sustainable Projects and Publications](#)
- 32. [Conversion of Primary Sources](#)

7.

Linguistics Meets Exact Sciences

Jan Hajič

Linguistics is a science that studies human language, both spoken and written. It studies the structure of language and its usage, the function of its elements, and its relation to other bordering sciences (psycholinguistics, sociolinguistics, etc.), of which the most important related field today is computer science. Diachronic linguistics studies language from the perspective of its development over time, whereas synchronic linguistics studies the function and structure of the current live language. General linguistics studies human language as a system, but particular languages (such as English, Mandarin, Tamil, or any other) are studied in detail as well. Due to its complexity, the study of language is often divided into several areas. Phonetics and phonology are related to speech (or more precisely, to the spoken language), whereas orthography deals with the standard written form of a particular language, including capitalization and hyphenation where appropriate. Morphology studies the composition of words by morphemes (prefixes, roots, suffixes, endings, segmentation in general, etc.) and its relation to syntax, which introduces structure into the description of language at the level of phrases, clauses, and sentences. Sentences are typically considered to be very important basic language units. Semantics and pragmatics study the relation of the lower levels to meaning and contents, respectively.

A description of a (correct) behavior of a particular language is typically called a grammar. A grammar usually generalizes: it describes the language structurally and in terms of broad categories, avoiding the listing of all possible words in all possible clauses (it is believed that languages are infinite, making such a listing impossible). An English grammar, for example, states that a sentence consisting of a single clause typically contains a subject, expressed by a noun phrase, and a verb phrase, expressed by a finite verb form as a minimum. A grammar refers to a lexicon (set of lexical units) containing word-specific information such as parts of speech (noun, verb, adjective, particle, etc.) or syntactic subcategorization (e.g., that the verb "to attach" has a subject and an indirect object with the preposition "to").

From the historic perspective, the scientific, economic, and political developments in the world before and during World War II were preparing the field of linguistics for what happened shortly thereafter. **Natural language** ☒ moved into the focus of people working in several scientific fields previously quite distant from linguistics (and other humanities as well): computer science, signal processing, and information theory, supported by mathematics and statistics. Today, we can say that one of

Ciberestructuras no transparentes, lenguajes de marcado, herramientas GIS y visualización de datos, plataformas de repositorios y análisis de datos, y los datos. enlazados **permiten actual escenario de las Humanidades Digitales.**

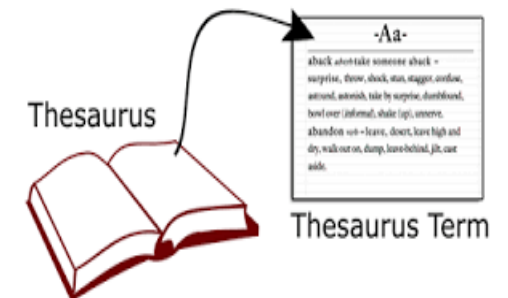
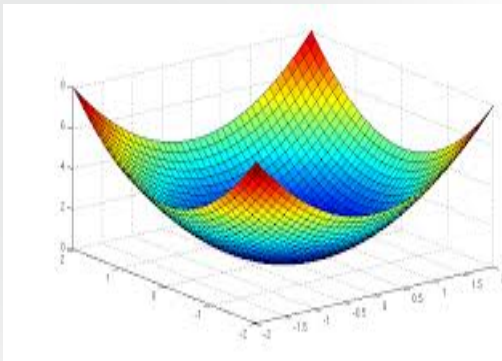
Sula, Chris Alen. Digital Humanities and Libraries: A Conceptual Model, 2012.



open source



Open Data



TaDiRAH - Taxonomía sobre Actividades de investigación digital en humanidades

Inicio Mi cuenta

Buscar

Búsqueda avanzada

Sobre...

A B C D E F G H I L M N O P R S T V

Análisis de clusters

El análisis de cluster es una forma de analizar los datos que clasifican un conjunto de información, en dos o más grupos mutuamente excluyentes basados en combinaciones de variables internas. El análisis de clústeres es útil para descubrir estructuras y patrones dentro de datos basados únicamente en una categoría seleccionada de similitud y diferencia. En la práctica, el análisis de cluster de un corpus de textos generalmente los agrupa de acuerdo a las similitudes y diferencias de las frecuencias de las palabras más frecuentes. Se ha demostrado que el análisis de clusters es altamente confiable en la atribución de autoría y en la identificación de género.

<http://www.vocabularyserver.com/tadirah/es/index.php>

Actividades de investigación

Términos específicos:

- ▶ [Meta-Actividades](#)
- ▶ [Captura](#)
- ▶ [Creación](#)
- ▶ [Enriquecimiento](#)
- ▶ [Análisis](#)
- ▶ [Interpretación](#)
- ▶ [Almacenamiento](#)
- ▶ [Difusión](#)

DIGITALE EDITION

- Convolut I
- Convolut II**
- Convolut III
- Convolut IV
- Convolut V
- Convolut VI
- Convolut VII
- Convolut VIII
- Convolut IX
- Convolut X
- Convolut XI
- Convolut XII
- Convolut XIII

SUCHE

Metadaten

Convolut II, Seite 5

Alte Zählung:
Conv. II, Halbbogen I, 1

Akademieausgabe:
XXI, 161,02-162,32

Auf dem Manuskript:
Neue Zählung: 5

Alte Zählung: 2. Conv. I, 1.

Signatur Kant: No. 1.

Andere Signatur: 2

Seite drucken

Fensteranzeige:

Linkes Fenster

Rechtes Fenster

Beide Fenster

Convolut II P_5

TRANSCRIPTION

2 Von den metaphysischen Anfangsgründen der Naturwissen-

3 schaft (*philosophia naturalis*) *1 geht die Tendenz <der Philosophie> auf die

4 Physik

5 Naturforschung (*Physica*) als einem System der empirischen

6 Naturkunde, als auf welche jene abzweckt; den was man die

7 (mit Newton) die mathematische Anfangsgründe der Naturwissenschaft nennt

8 ist zwar auch die *2 gleichfalls auf Principien *a priori* ist aber nicht

9 ein Erkenntnis aus Begriffen sondern aus der Construction derselben

10 (den jenen correspondirenden reinen Anschauungen) und macht keinen

11 Theil der Naturphilosophie aus sondern ist *3 hiebey ~~er~~ ~~verwischet~~ nur ein

12 ~~er~~ ~~wie~~ ~~wohl~~ vortreffliches ja unentbehrliches Instrument zu dieser

13 ihrem Zwecke zu gelangen; ~~er~~ ~~welche~~ ~~welcher~~ die empirische Naturforschung

14 (*Physica*) ist.

15 Es ist aber unmöglich aus bloß empirischen Erkenntniß Begriffen

16 ein System zu zimmern. Es wird jederzeit ein ~~er~~ ~~zusamen~~ gestoppertes ~~er~~

17 zusamengestoppertes

18 Aggregat #bleiben von Beobachtungen dieser oder jener Eigenschaft der

19 Materie bleiben was zwar ansehnlich aber doch immer nur fragmen-

20 tarisch wachsen kañ und *5 we man still stehen kañ wo man will

21 weil es an der Idee fehlt mangelt welche ein innerlich #begründetes

22 und zugleich sich selbst #begrenzendes Gantze ausmacht; ein solches aber

23 also kañ nicht anders als nach Begriffen *a priori* zu Stande kom-

24 men.

25 Gleichwohl verlangt der philosophische Naturforscher daß dieser Schritt

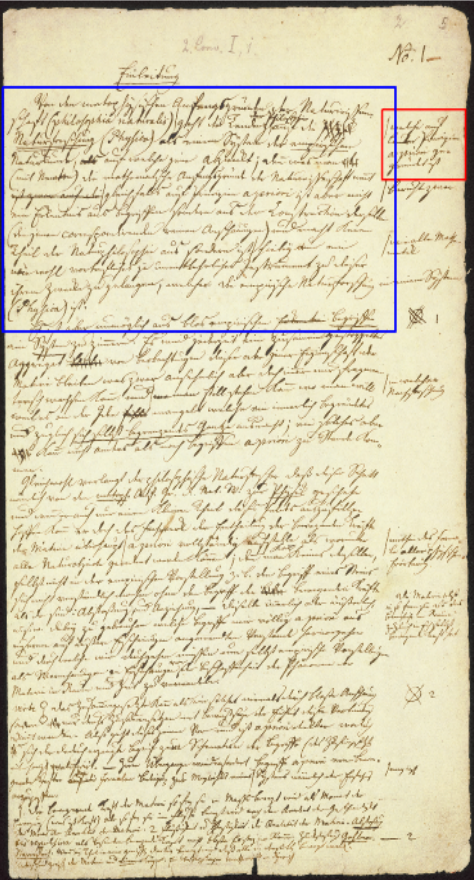
26 nämlich von den metaph. Anf. Gr. d. Nat. W. zur Physik geschehe

27 und weñ er auch nur einen kleinen Theil dieses Feldes auszufüllen

28 hoffen kañ er doch das gründliche Erkerwerk der Eintheilung der

Convolut II P_5

TRANSCRIPTION FAKSIMILE



THE ATLAS *of* EARLY PRINTING

| [About the Atlas](#) | [15th Century Printing](#) | [The 15th Century Book](#) | [Printing Press Model](#) | [Sources](#) | [Links](#) |

1450 1455 1460 1465 1470 1475 1480 1485 1490

Use the slider to change the date

1450 - 1500

Select a layer to display on the map

Spread of Printing

Output by Location

Universities

Show pre-1450 Locations

Paper Mills

Hide pre-1450 Locations

Fairs

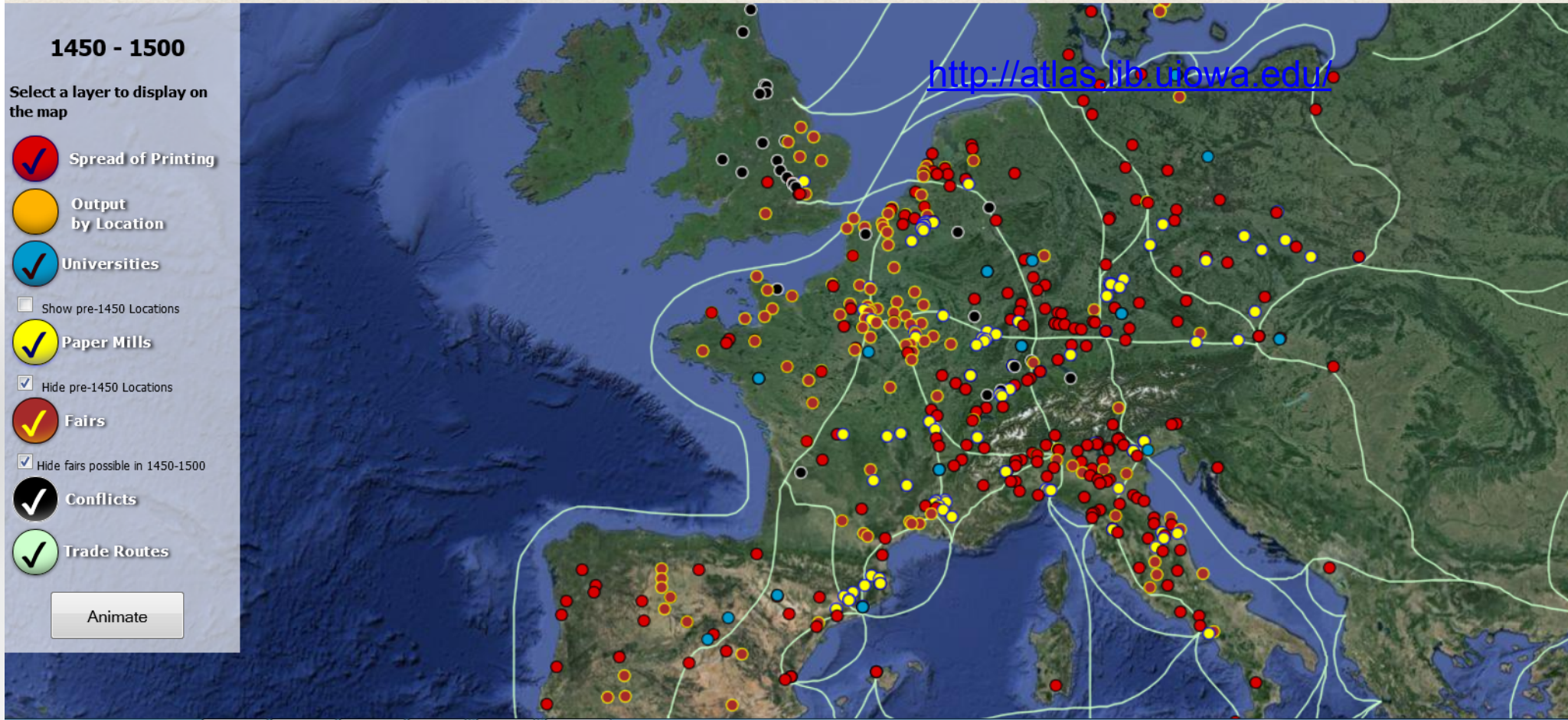
Hide fairs possible in 1450-1500

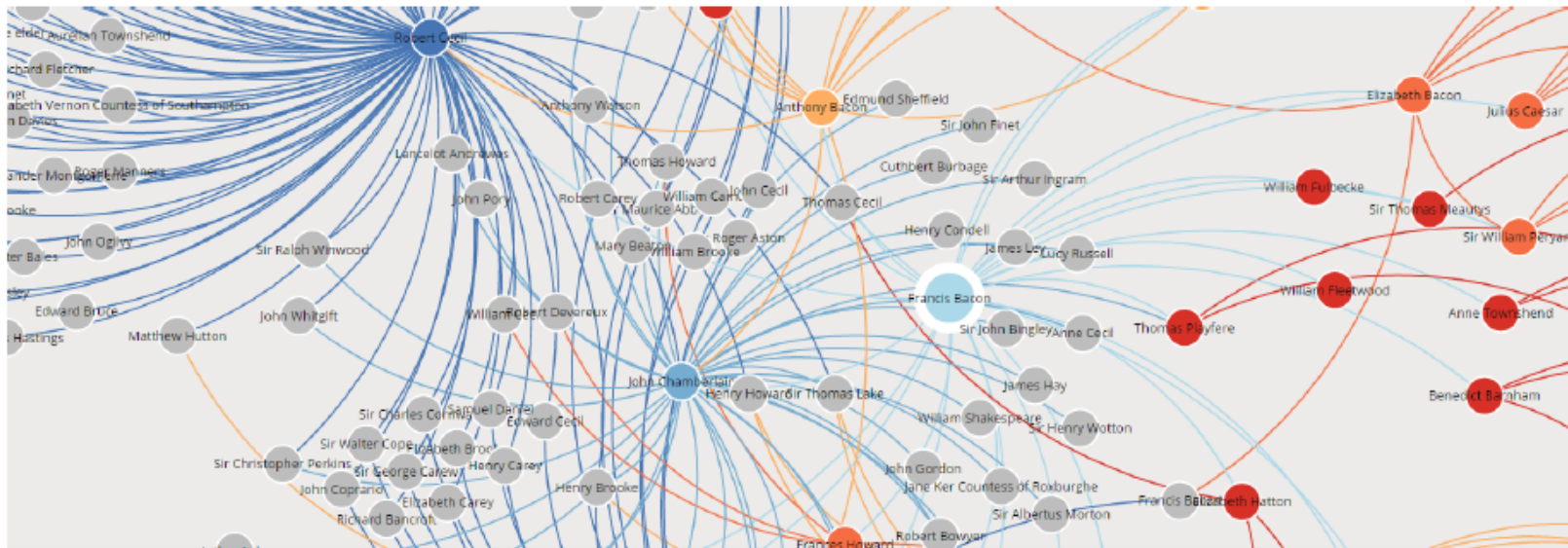
Conflicts

Trade Routes

Animate

<http://atlas.lib.uiowa.edu/>

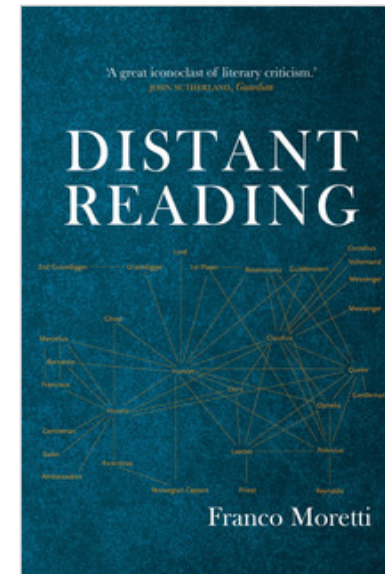
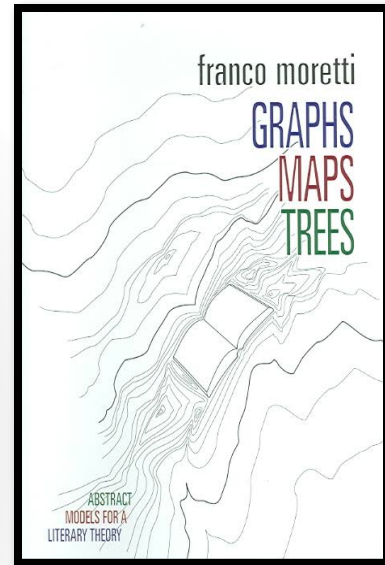




Six Degrees of Francis Bacon <http://www.sixdegreesoffrancisbacon.com>

Distant Reading vs. Close Reading

Franco Moretti, 2005 - 2013



ADSO *Análisis distante del soneto castellano de los Siglos de Oro*

Proyecto ADSO ▾

Corpus de métrica ▾

Sistema de escansión ▾

Español ▾

Proyecto ADSO

El objetivo del proyecto ADSO es realizar un análisis **macroanalítico y distante** del **soneto castellano de los Siglos de Oro**, desde sus primeras manifestaciones renacentistas (Garcilaso de la Vega) hasta los últimos poemas del Barroco (Sor Juana Inés de la Cruz). Para ello se aplican **métodos computacionales** con el fin de detectar sus rasgos generales recurrentes, tanto en aspectos semánticos como métricos.

A diferencia de estudios anteriores, el objetivo no es **analizar** los rasgos específicos de los sonetos de unos pocos autores canónicos, sino **los rasgos comunes de todos los sonetos escritos en castellano durante los siglos XVI y XVII mediante la aplicación de técnicas computacionales**. Como resultado del proyecto se espera poder especificar cuáles son esos rasgos literarios comunes a todo el periodo tanto **métricos** como **semánticos**

Buscar ...

Entradas recientes

Humanidades Digitales Hispánicas
2017

Guía de anotación actualizada

Publicación del modelo automático
de escansión

Los valencianos Cristóbal de Virués y
Jerónima Agustina Benita

Convocatoria de dos plazas
temporales de técnico superior para
colaborar en el proyecto ADSO

Corpus de sonetos castellanos del Siglo de Oro

Proyecto financiado por la [Fundación BBVA](#): ayudas a equipos de investigación científica 2016. Proyecto "[Análisis distante de base computacional del soneto castellano del Siglo de Oro \(ADSO\)](#)"

Introducción

Corpus formado por sonetos escritos en castellano entre los siglos XVI y XVII.

Cada soneto ha sido anotado en XML siguiendo el estándar [TEI](#). Además de encabezado e información estructural, cada soneto tiene representado formalmente los **patrones métricos** de cada verso.

Anotación

La anotación de los patrones métricos se ha realizado de manera semi-automática. Primero se han procesado todos los sonetos con un sistema automático de escansión métrica, que asigna a cada verso su patrón métrico. Después se ha validado a mano cada patrón métrico y se han corregido los errores detectados.

Langas

- lenguas generales coloniales
- quechua, guaraní, tupí y aimara
- textos políticos/administrativos siglos XVI-XIX
- www.langas.cnrs.fr

LANGAS

Lenguas generales de América del Sur

Type to enter a caption.

Título			Autor	Fecha
Proclama al Pueblo Paraguayo			Belgrano, Manuel	1810
#				
1	Porômô rându	Poromorandu	Proclama	Proclama
2	Paraguaỹ gua reta, heco aguĩ yeibe bae, haè heco catu pĩrĩbe bae :	Paraguaygua reta, heko aguyjeive va'e, ha heko katupyry va'e :	Nobles, fieles y leales Paraguayos :	Paraguayos cuyo modo de ser es noble y valiente
3	Ayquaa uca peme ayu haba Exma Junta recobia râmô,	aikuaa uka pe(ẽ)me aju háva Excelentísima Junta rekovia ramo,	Vengo de Representante de la Exima Junta Provisional Gubernativa	os hago saber que vengo como representante de la Excma Junta
4	ñande Rey D.n Fern.do e 7° rera pĩpe, co ara ñande yoquaibae	ñande Rey Don Fernando 7° réra pýpe, ko ára ñande jokuái va'e	que à nombre de S.M. el Sor D.n Fernando 7° rige las Provincias del Rio dela Plata,	que a nombre de nuestro Rey Don Fernando 7° ahora nos gobierna
5	haè aikuaa ucabe peme, ayuha Capayu eta rubicha râmô,	ha'e aikuaa ukave pe(ẽ)me, ajuha Kapaju eta ruvicha ramo,	y de General en Gefe del Exto,	y os hago saber que vengo como jefe de numerosos soldados
6	arubaè pẽ ndepĩbo,	aru va'e pende pyvo,	q.e à vista de vuestros clamores, ha dispuesto	a los que traigo junto a vosotros

Proclama al Pueblo Paraguayo, versión paleográfica y moderna en guaraní y español

- 117 documentos
- transcripción + transliteración + traducción (+ transliteración español original)
- primeros escritos
 - variación ortográfica:
llaqta ~ llakta ~ llacta ~ llagta
 - segmentación de palabras:
llactacunapi - **lacta**ma - **llaqta**masinkunamanmi

á â ã ä é ê ë ě ê í î ï ð ó ô õ ö
ú û ü û ŷ Ÿ ÿ ŷ ŷ ŷ š š č č ç ñ ’

En posición

Indiferente

En versión

amerindia

Corpus

Guaraní

Quechua

Tupi

Búsqueda avanzada



Buscar en paleografía transliteración

Buscador polígrafo

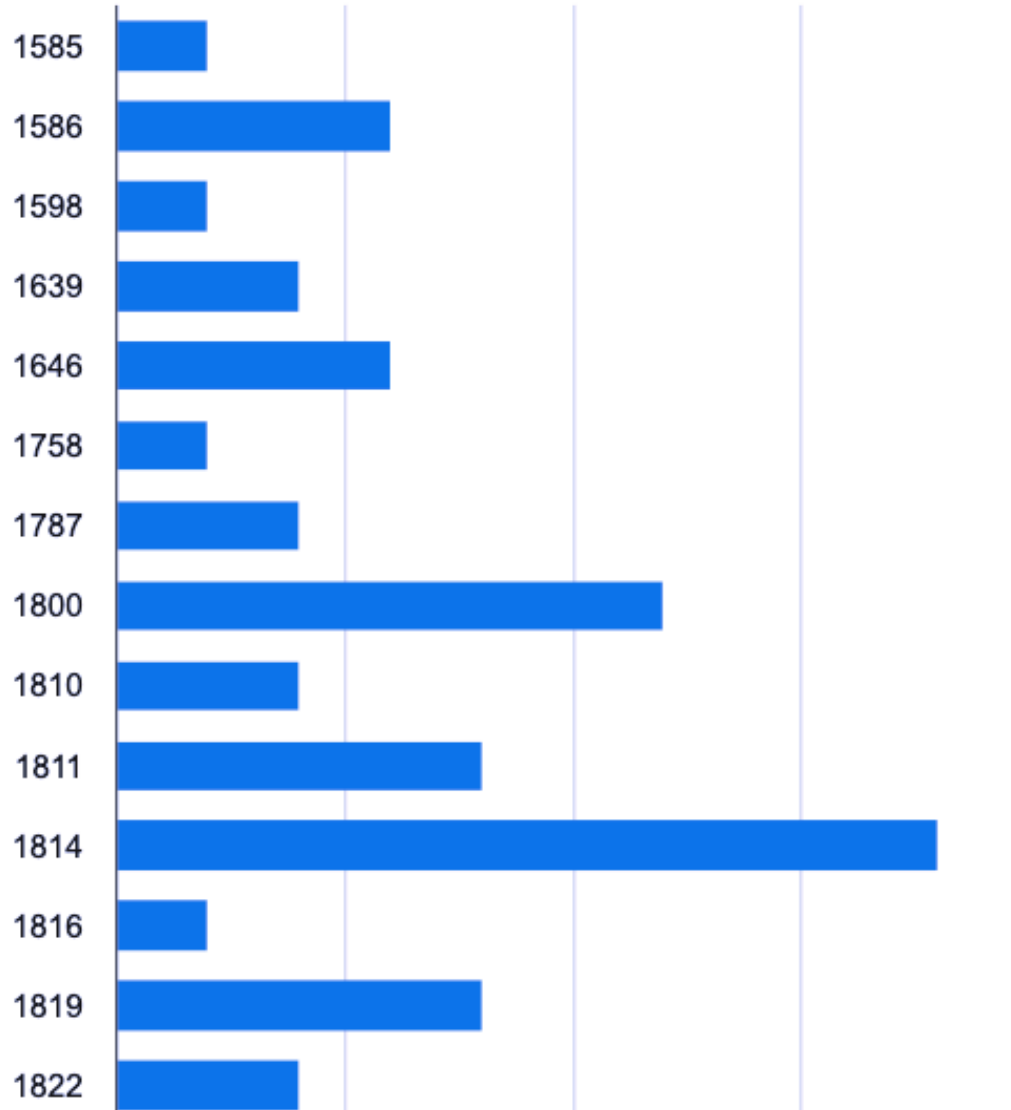
Llipinmi kay Ilaqta manta runa o Ilipillanmi rin. Todos los indios de este pueblo se han huido y le dejan limpio de gente.	Vocabulario de la lengua general de todo el Perú llamada lengua Qquichua, o del Inca – A-LL	1608
Auqui-Camachec Kcollana mama patriac runancunata, pata llactacunapi causac llactama sincunamanmi-nin	Proclama	1822
Apu Don Juan Josef Castelli Apucunat Buenos Airespi Tantascacunac Lantim: Cai Perú Llacta Runacunamam.	Proclama de Castelli en Charcas	1811
Qankunaga upya-machay chakra kinraypi saksapukuyllakta huchaptiykihinaña mana confesasaq wañun.	Cartas de Cotahuasi	1616
inti killakta quyllurkunaktapaș	Symbolo Feria Segunda	1598

Búsqueda de "Ilaqta"

- equivalencias fonema/grafema
- fonetización del corpus textual
- fonetización de la búsqueda
- búsqueda en textos fonologizados
- correspondencia con originales

RÉALISATIONS PHONÉTIQUES	GRAPHIE ACTUELLE	GRAPHIES ANCIENNES PROTOTYPIQUES	PALÉOGRAPHIQUES
[i]	-i- -í-	-i- (<i>rupi</i>) -í- (<i>cunumî</i>) -y- (<i>ychupe</i>) -î- (<i>îru, mî</i>)	-ì- (<i>oroì</i>) -ý- (<i>mondĭý</i>) -ÿ-
[j] (semi-cons.)	-j-	-y- (<i>yaiquaa</i>)	
[i_] (semi-voc.)	-i-	-y- (<i>aypo</i>) -i- (<i>acoi</i>) -ĩ- -ÿ- (<i>ruguâÿ</i>)	-ĩ- (<i>rũĩ</i>) -î- (<i>oñoîrârô</i>) -î- (<i>acoî</i>)
[ĩ] contxt nasal		-î- (<i>oquirîrî</i>)	
[j] V_V o ataque	-j-	-y- (<i>aguïye</i>)	-i- (<i>aguïebe</i>)

Equivalencias /i/



Resultados por
año para
"libertad"

Recogito

- Plataforma de anotación semántica
- Anotación de textos e imágenes
- Antigüedad y modernidad
- recogito.pelagios.org

Recogito



AN INITIATIVE OF
Pelagios
commons

Type to enter a caption.

Capítulo I

¿Quién fue el primer descubridor de estas provincias del Río de la Plata?

2

Después que el Adelantado Pedro de Vera, mi rebisabuelo, por orden de los Reyes Católicos don Fernando y doña Isabel, conquistó Gran Canaria, que antiguamente se dijeron Fortunadas, luego el Rey de Portugal mandó poblar las islas de Cabo Verde, que están en la equinoccial, y cursar el comercio de las minas de Guinea, y por el consiguiente el año de 1493 salió de Lisboa un capitán llamado por orden del mismo Rey don Juan, a hacer navegación al Occidente, al mismo tiempo que Cristóbal Colón volvió a España del Indias. Este capitán Américo llegó a Cabo Verde, y continuando su jornada pasó la equinoccial de este cabo del Polo Antártico hacia el Mediodía, de manera que llegó a reconocer la tierra y costa del Brasil junto al Cabo de San Agustín, que está ocho grados de la parte del Norte, donde, corriendo aquella costa, descubrió muchos puertos y ríos caudalosos, y toda ella muy poblada de gentes caribes y caribes septentrionales se llaman Tobaiaras y Tamoios. Los australes se dicen Tupinambás y Tupinás; son muy belicosos, y hablan todos lenguas diferentes, aunque con alguna diferencia: andan todos desnudos, en especial los varones, así por el calor de la tierra, como por ser antigua costumbre.

kings of

rey

King of

3

2

2

2

2

2

2

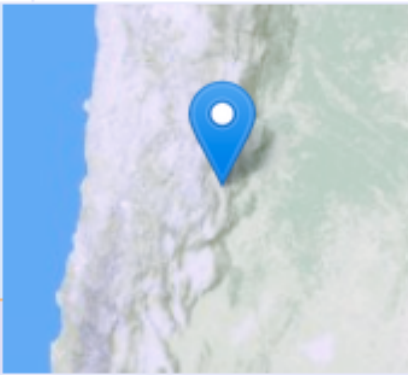
atraviesa toda la gobernación del Tucumán, y nace de las Cordiller
de Santiago del Estero regando muchas tierras y pueblos de indios

st
cu
no
n
gu

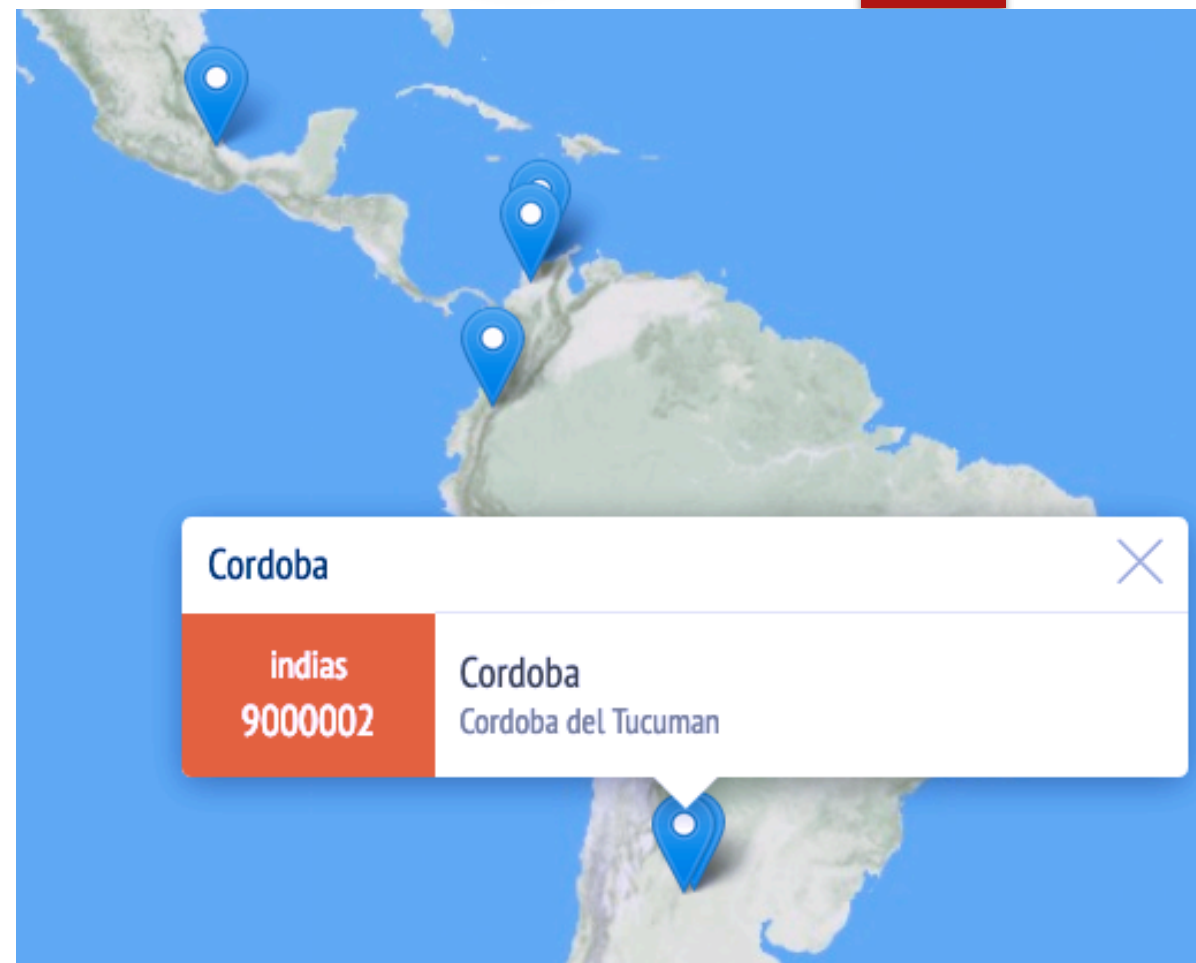
📍+ Place 👤+ Person ★+ Event

Santiago
indias:9000017
Santiago del Estero, El Barco

⚠️ Automatic Match Confirm Change 🗑️



Sugerencias automáticas para
anotación de lugares



Cordoba

indias
9000002

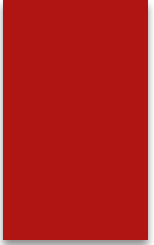
Cordoba
Cordoba del Tucuman

Distintos lugares, mismo nombre

LatAm gazetteer

00130;1003944;Aconchi;Pueblo;Pueblo;Sonora;PRGDSS00
00131;;Aconichi;Pueblos y tierras indígenas;Pueblo;;
00132;;Aconichi;Isla;Isla;;
00133;;Acokquiya;Montaña;Cerro;Tucuman;PRRPTU00
00134;6000338;Acopia;Pueblo;Pueblo;Corregimiento de Qu:
00135;8000200;Acora;Pueblo;Pueblo;Chucuito;JUCCPUCH
00136;;Acori;Rio;Rio;;
00137;6000583;Acoria;Pueblo;Pueblo;Corregimiento de Ang
00138;6000339;Acos;Pueblo;Pueblo;Perú;SUPE0000
00139;;Acos;Provincia;Provincia;Corregimiento de Quisp:
00140;;Acosta;Pueblo;Pueblo;Brazil;
00141;6000625;Acostamba;Pueblo;Pueblo;Perú;SUPE0000
00142;;Acostambo;Pueblo;Pueblo;Corregimiento de Huanta;
00143;6000805;Acotama;Pueblo;Pueblo;Corregimiento de Cl

Información geográfica estructurada



le de Pu- el N hay laman de nde tam- palmente Ciudad de r ser allí lera , so- r en los y Marzo. esta Pro- 3y almas. an Felipe nombre

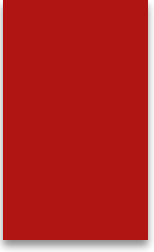
ACONCHI, Pueblo de la Provincia y Gobierno de Sonora en Nueva España.

ACONICHI, Pueblo de Indios de la Carolina Septentrional, situado á la orilla del rio Eno.

Tiene el mismo nombre una Isla que hay en medio del rio Dan, en la propia Provincia.

ACONQUIJA, Cerro de la Provincia y Gobierno del Tucuman, en el distrito de la Ciudad de Catamarca, y

Diccionario Histórico Geográfico de Antonio de Alcedo (1786)

- 
- variación ortográfica (j/g/x)
 - sinónimos (ej: montes, montaña, morro, sierra)
 - anáforas ('Tiene el mismo nombre', 'Otro pueblo')
 - ambigüedades (ej: 'Chorillos')
 - secciones especiales ('Correcciones', 'Vocabulario', etc)
-

- 5 volúmenes, ~990k palabras, ~18k entradas
- 35% son accidentes geográficos (río, valle, etc)
- ~2k registros nuevos para Recogito
- variantes ortográficas aún no exploradas!



Type to enter a caption.

JOIN US!



hdcaicyt.github.io




hdcaicyt@gmail.com



[@HDCAICYT](https://twitter.com/HDCAICYT)





Revista de
Humanidades
Digitales



<http://revistas.uned.es/index.php/RHD/index>

Asociación Argentina de Humanidades Digitales



Sitio web: <http://aahd.net.ar/>

Facebook: <https://www.facebook.com/aahdArg/>

Grupo cerrado en Facebook:

<https://www.facebook.com/groups/humanidadesdigitalesargentina/>

Twitter: [@aahdArg](https://twitter.com/aahdArg)

Repositorio AAHD en Acta Académica: <https://www.aacademica.org/aahd>

Mailing list de la AAHD: <http://listserv.uleth.ca/mailman/listinfo/aahd-l>