

Filología digital: transcripción automática de mss, colación asistida y creación de corpus textuales

Mónica Castillo Lluch
Andrea Escobar Castillo
José Manuel Frajedas Rueda
Francisco Gago Jover

Curso de verano, 2024

UIMP Universidad Internacional
Menéndez Pelayo

Introducción

Mónica Castillo Lluch

Selección de artículos de Mónica Castillo Lluch sobre el Fuero Juzgo (y otros fueros) (2012-2023)

Mónica Castillo Lluch, *Tel fils, tel père : Ferdinand III dans le processus de planification du castillan (étude linguistique du Fuero juzgo)*, 2011. https://serval.unil.ch/fr/notice/serval:BIB_5E3C78DF48FB

Mónica Castillo Lluch, «Las lenguas del *Fuero juzgo* : avatares históricos e historiográficos de las versiones romances de la Ley visigótica (I)», *e-Spania* (Revue interdisciplinaire d'études hispaniques médiévales et modernes), 13, Juin 2012. <https://journals.openedition.org/e-spania/20994>

Mónica Castillo Lluch, «Le *fuero juzgo* dans l'histoire de la langue espagnole», *e-LHR (e-Legal History Review)*, 14, Juin 2012 <http://www.iustel.com/v2/revistas/detalle_revista.asp?id_noticia=412118&d=1

Mónica Castillo Lluch / Ricardo Pichel, «El códice López Ferreiro del *Fuero juzgo*: estudio histórico, filológico y lingüístico», *Revue de linguistique romane*, tome 79, n°. 313-314, 2015, 123-168.

Mónica Castillo Lluch, «El orden de palabras en los fueros castellanos del siglo XIII», in López Izquierdo, Marta / Castillo Lluch, Mónica (eds.): *El orden de palabras en la historia del español y otras lenguas iberorromances*, Madrid, Visor (Visor Lingüística), 2015, 279-318.

Mónica Castillo Lluch, «Las fechas del *Fuero juzgo* : avatares históricos e historiográficos de la versión romance de la ley visigótica (II)», in Narbona Jiménez, Antonio et Araceli López Serena (coords.): *El español a través del tiempo. Estudios de lingüística histórica ofrecidos a Rafael Cano Aguilar*, Sevilla, Editorial de la Universidad de Sevilla, 2016, 47-68.

Mónica Castillo Lluch, «La historia de la lengua española vuelve por sus fueros», in María Luisa Arnal Purroy et al. (ed.) : *Actas del X Congreso Internacional de Historia de la Lengua Española*, Zaragoza : Cometa, 2018, I, p. 129-155.

Mónica Castillo Lluch, « Los fueros en la obra de Ramón Menéndez Pidal», in Inés Fernández-Ordóñez (ed.) (2020), *El legado de Ramón Menéndez Pidal (1869-1968) a principios del siglo XXI*, vol. II, Madrid, CSIC, Anejo de la *Revista de Filología Española*, 81-105.

Mónica Castillo Lluch / Charles Mabille «El Fuero Juzgo en el ms. BNE 683 (1755) de Andrés Marcos Burriel», *Scriptum digital* 10, 2021, 75-107.

Mónica Castillo Lluch / Andrea Escobar Castillo, «El fuero de Medinaceli concedido a Murillo el Fruto: nueva edición y estudio lingüístico», *Revista de historia de la lengua española*, 2022, 17, 113-143.
<https://doi.org/10.54166/rhle.2022.17.05>

Mónica Castillo Lluch «Fuero parece, Real lo es: BH MSS 345 de la Biblioteca Histórica “Marqués de Valdecilla”» *post* de PhiloBiblon 2023 n. 2 (marzo)

<https://update.lib.berkeley.edu/2023/03/18/fuero-parece-real-lo-es-bh-mss-345-de-la-biblioteca-historica-marques-de-valdecilla/>

Mónica Castillo Lluch / Charles Mabille, «Hacia un *stemma codicum* del Fuero Juzgo desde el Humanismo hasta hoy», in Junquera Martínez, Alejandro, Egido Fernández, María Cristina, Morala Rodríguez, José Ramón (eds.), *Actas del XII Congreso Internacional de Historia de la Lengua Española : [León, 16 al 19 de mayo de 2022]*, León, Universidad, 2024, 539-555 <https://buleria.unileon.es/handle/10612/21710>

Mónica Castillo Lluch / Carmen del Camino Martínez (2024) « Los manuscritos del Fuero Juzgo: denominaciones antiguas y signaturas modernas » *post* de PhiloBiblon 2024 n. 4 (mayo)
<https://update.lib.berkeley.edu/2024/05/08/mssfuerojuzgo/>

Mónica Castillo Lluch / Yolanda García López (en prensa) «The Visigothic Code and the *Fuero Juzgo* : The Transmission and Translation of Law from Latin to Romance», Noel Lenski / Damián Fernández (eds.) *Lex Visigothorum*, Cambridge University Press.

o pequeño o grande que fuer p^{uado} q
fisiere este peccado: matenente el príncipe
o el alalde los manda castigar: i aun solo
to aya aquella pena la qual dieron los sa-

cerdotes en su decreto. el tercero anno de

mi regno por tal peccado. El Rey. do flia

uno rescindo. ¶ Delos que yaze con las

mujeres. delos padres. & delos h^{er}manos.

En la ley desuso auemos dicho q

la pena deuen auer los que casan co

las parientas. mas toda uja por que no

deuen auer menor pena aquellos q yazi

con las mugieres delos padres o delas h^{er}ma-

nios. mandamos en esta ley que neq^uun

ome non se iazer con la barragana del

padre o de su h^{er}mano. o con la que sopiere

con quie yogo su padre. o su h^{er}mano alg

una uez. si quier sea libre si quier sieru.

nin el padre non aya que ueer o la mu-

gier que yogo el fijo. Et si alguno fizi

en tal cosa sabiendo: su buena ayan-

toda sus hijos legitimos. si los ouiere. i

si no los ouiere: ayanlo. sua herederos

mas. primos. y el sea echado dela tierra

por pena por siempre i este en poder dal-

guin omne quel laga sollar pena regalo

por su vida. ¶ Delos parientes delos

padres. i delos des^us padres.

En la mugier se parte del mando o

dicho o con uera. Que los ca-

sados non se puedan partir. Que los es-

padas non se partan. ¶ Si la quis-

que se parte del marido n^oechas que

mu^g que fue dunda de su marido

nenguno non se case con ella si no sopie-
re que la devo atra m^{er}ge por escrito o
por testimonia. si lo fizieren el sacerdote dela
abad o el vicario o el alalde depuca q
lo sopiere. si fuer p^{sona} que no pue-
da constar que se departan: fagalo sabi-

al Rey. Et si fueren plomas de menor

fa: fagalo: parte luggo. asⁱ que la mugier

que caso anna la uoluntad del príncipe mi-

ndo en adulterio. i a quel que la come p

mu^g. Sean menuda en poder del príncipe;

massto. i fagadellos lo que estricto. Toda

mia en tal maniera si non en juzgado por

juzgio. o si el marido deula mu^g o deue-

ro: deuela mu^g auer laq^{ue} am^{er} quel die-

ni. y el non deue auer nada de las cosa

de la mugier. Et si alguna cosa le ame-

tomado. o en auendido: todo lo entaguez-

ala mugier: i si la mugier seiendo en;

poder del mundo por engaño o por arte le-

dien la quarta parte dun dino al mari-

do quel devo. nol uala. am^{er} que gelo de

por escrito. mas q^{uo}d dia la mugier:

por aquell escrito: tod deue ouir a el

la. ¶ Que los casados non

se pudan partir. ii. Ley.

En pecando ea yager o la mugier

aienas: mayor pecando ea deuen la

uya con que se casó: por su grido i por q

son algunas que por cobardia o por lau-

deria las sig^{ue} mugieres. i aun casar o

laienas. faginen esta constancia q

nengun omne no tere su mugier si non

por adulterio. nin se para della por escr

Enregistrement automatique paleográfica 2_7_23_ Enregistré dans mon Mac

Accueil Insertion Dessin Conception Mise en page Références Publipostage Dites-le-nous Commentaires Modification Partager

Cambria (C... 12 A Aa Aa) Styles Volet Styles Dicter Rédacteur Crée un PDF et Demander des signatures

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18

{4}to aya aquella pena la qual dieron los sa
{5}cerdotes en su decreto. el tercero anno de
{6}n^ouestro regno por tal peccado. El Rey. do<n> fla
{7}uo rescindo. ¶ Delos que yaze<n> con las
{8}mujeres. delos padres. & delos h<er>manos.
{9}EN la ley desuso auemos dicho q<ua>l
{10}pena deuen auer los que casan co<n>
{11}las parientas. mas toda uja por que no<n>
{12}deuen auer menor pena aquellos q<ue> yaze<n>
{13}con las mugieres delos padres o delos h<er>ma-
{14}nos. mandamos²⁶⁶ en esta ley que ne<n>gun
{15}om<n>e non ose iazer con la barragana del
{16}padre o de su h<er>mano. o con la que sopiere
{17}con quie<n> yogo su padre. o su h<er>mano alg
{18}una uez. si quier sea libre si quier sieru<a>.
{19}nin el padre non aya que ueer²⁶⁷ co<n> la mu
{20}gier que yogo el fijo. Et si alguno fizi
{21}ere tal cosa sabiendo: su buena ayan-
{22}toda sus hijos legitimos. si los ouiere. &

266 En Murcia (RAE 63 b: *ennademos*) La coincidencia es con más: S.B., Malp. 2, Camp. Esc. 1, 2 y 4, Bex.
267 yaga en resto (RAE 63b)

Page 154 sur 214 72257 mots Espagnol (Espagne) Accessibilité : consultez nos recommandations Focus 169%



Texto del Fuero Juzgo (654, después 1241)

Tradición manuscrita compleja (texto conservado en 50 mss)

Por hacer:

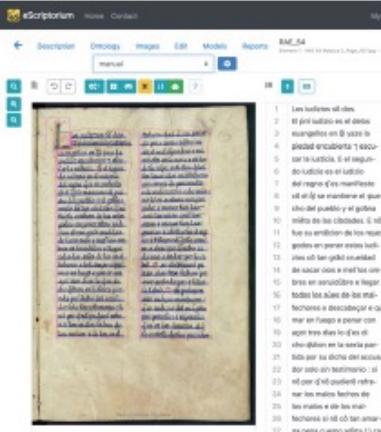
- transcripciones (solo 5 mss editados)
- comparaciones de las diferentes versiones para su estudio

Proyecto piloto: transcripción automática del ms. RAE 54
con eScriptorium

FUERO JUZGO digital

dirección: Mónica Castillo Lluch
apoyo informático: Andrea Escobar Castillo
asistentes: Orléane Chioccola y Domitille Gaillard

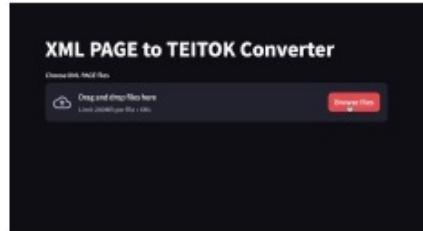
1 Uso de eScriptorium para transcripción semiautomática



2 Corrección humana



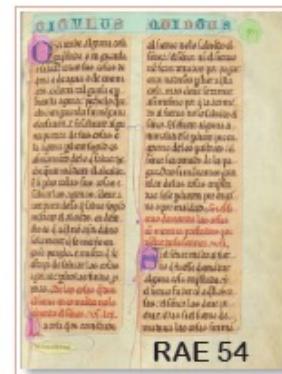
3 Volcado en TEI:TOK (script)



4 Colación asistida para comparar versiones (CollateX-Notebook)



El proyecto Fuero Juzgo digital pretende superar la falta de conocimiento de este texto importante para la historia de la lengua y del derecho, mediante la producción y puesta en línea de ediciones de sus testimonios antiguos más interesantes. La comparación de los testimonios permitirá explorar los contrastes entre las dos recensiones, fernandina y alfonso, y mostrará las diferentes formas de la recensión alfonso, conservadas en manuscritos de los siglos XIII al XV.



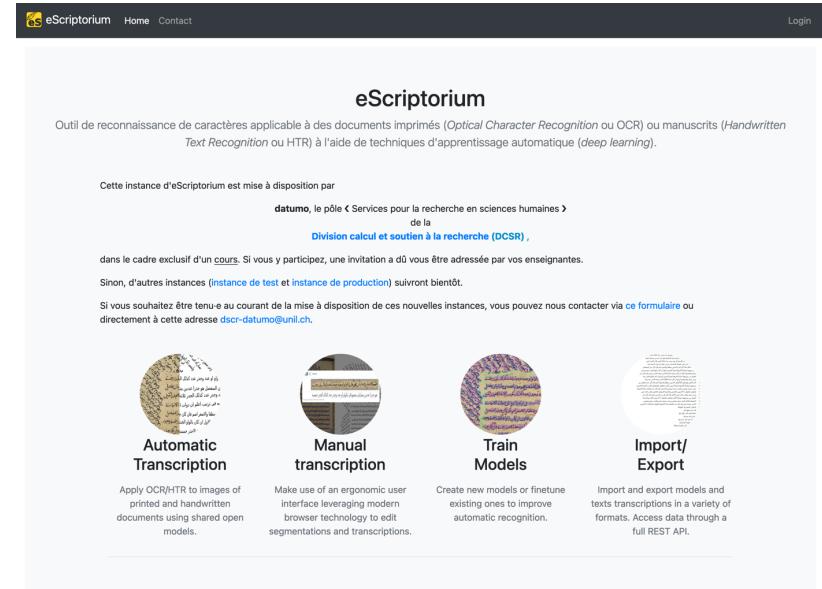
RAE 54

En junio de 2024 hemos finalizado la transcripción de RAE 54 (ca. 1260, recensión alfonso) con el programa eScriptorium y revisión humana. Su subida a TEI:TOK está prevista antes de finales de este año.



- 1 C/S de MA (2023A)
- 6 estudiantes
- 2 AD perfil HD
- 1 experto ENS/École nationale des chartes
- Financiación Faculté des lettres UNIL → ingenieras pedagógicas + Centro de cálcul
- Instalación de eScriptorium en un servidor de la UNIL → proyecto piloto para la Facultad de letras UNIL

<https://escriptor.dcsr.unil.ch/>

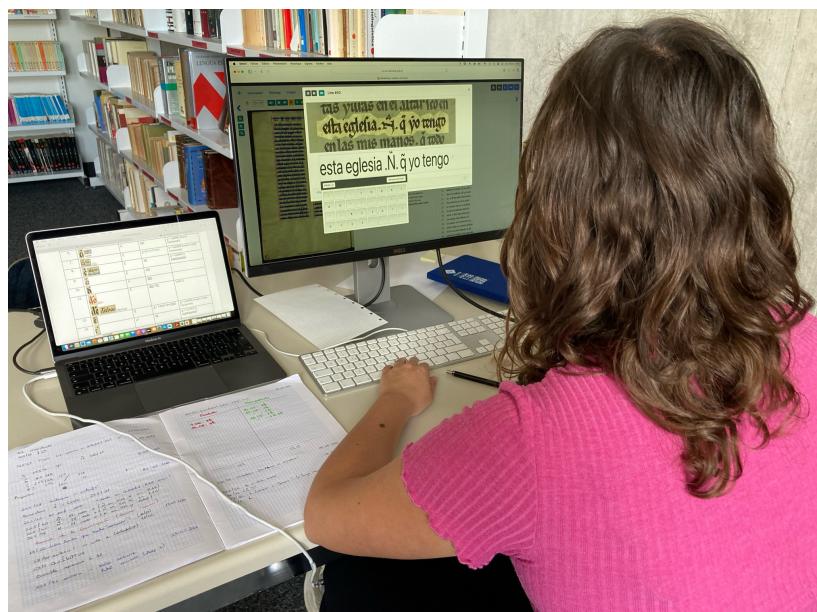


Objetivos 2023-2024 (transcripción RAE54)

etapa 1 (C/S 2023A) : 60 folios corregidos manualmente (120 imágenes) por los 6 estudiantes



etapa 2 (2024P) : revisión de los 60 folios + entrenamiento de un modelo para la transcripción optimizada de los 172 folios restantes por une équipe de 2 asistentes estudiantes, Andrea Escobar y Mónica Castillo



Andrea Escobar Castillo

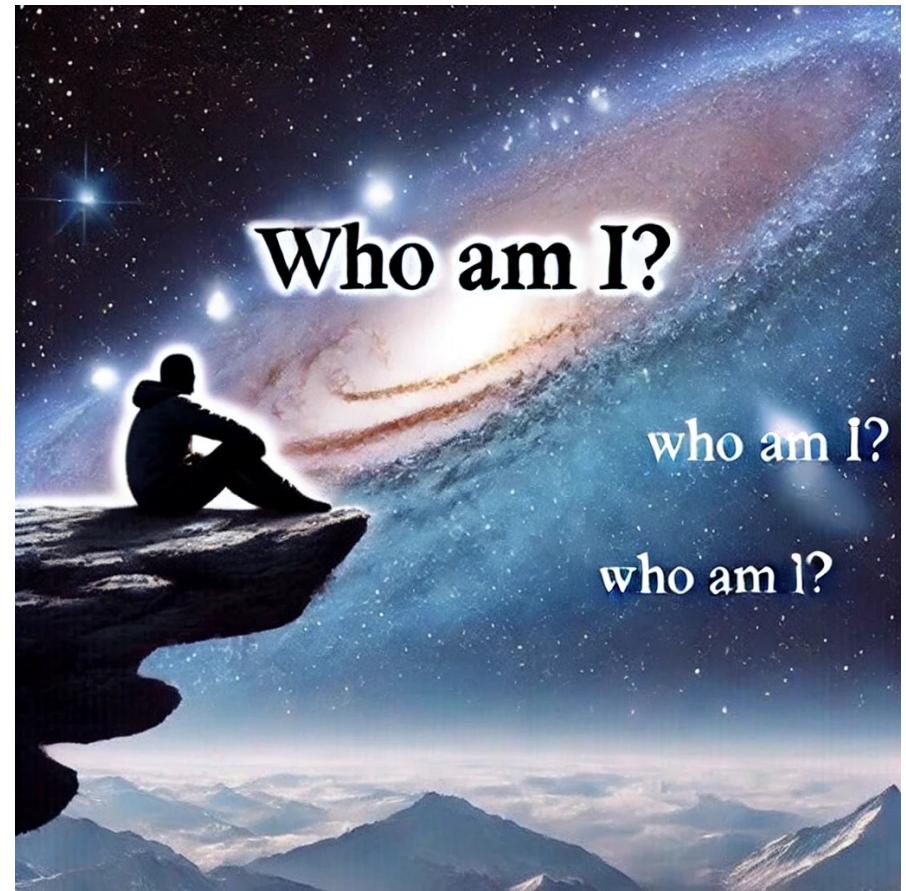
>-

~ % whoami

MA en español, informática para las ciencias humanas y especialización en historia del libro y edición crítica de textos

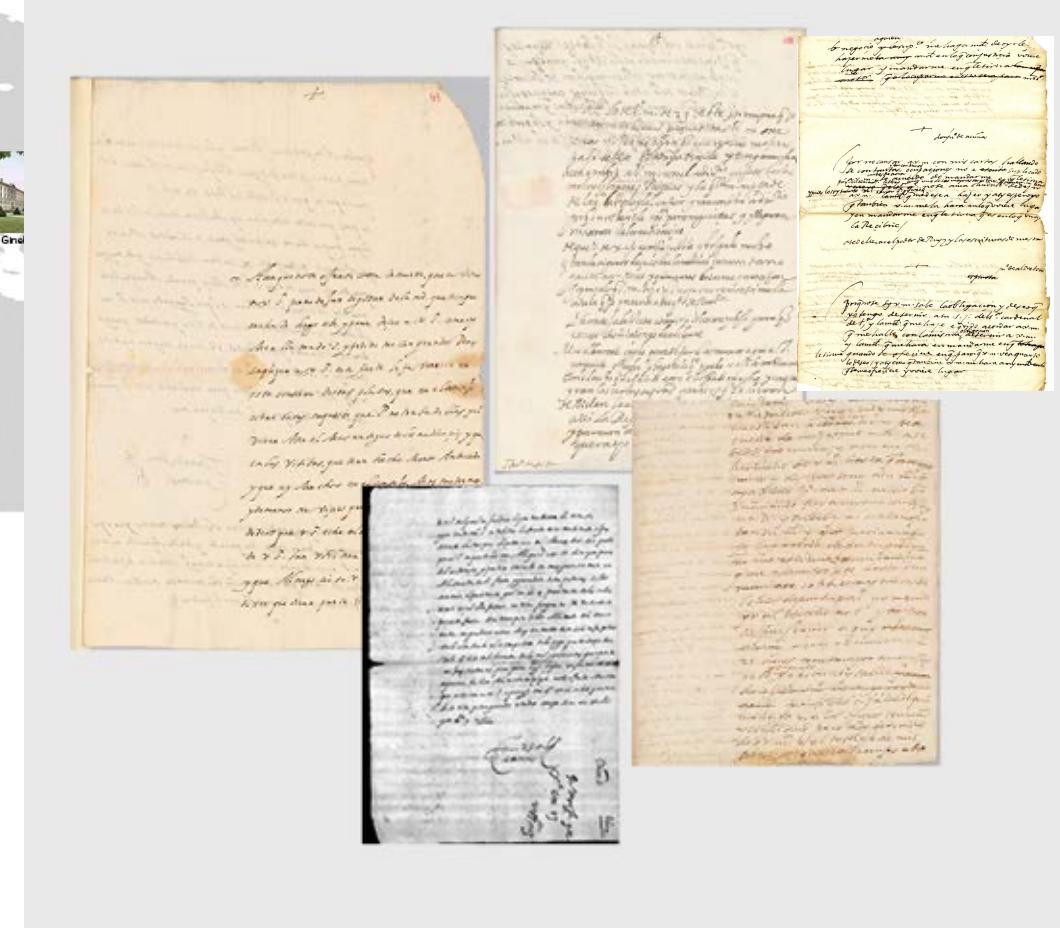
Tesis

- Caracterización lingüística (estudio idiolectal) de un miembro de la nobleza madrileña de las cortes de Felipe II y Felipe III
- Comprensión de la escritura por delegación gráfica
- Empleo de diferentes métodos informáticos desde la adquisición de los datos hasta su explotación (transcripción automática, anotación gramatical, estilometría)



Corpus

- Documentos: 335 documentos (+/- 3-8 páginas por documento)
- Fragmentación del antiguo fondo de la casa de Altamira
- Acopio de datos:
 - Reproducciones: distinta calidad de digitalización
 - Amplia variación gráfica
 - Pobre desempeño de modelos de transcripción disponible



Presentación de los participantes al curso

1. ¿quiénes son?
2. ¿de dónde vienen y adónde van?
3. ¿por qué se han inscrito al curso?
4. ¿con qué perspectivas de aplicación en qué proyectos personales?
5. ¿qué conocimientos tienen de los contenidos del curso?
6. ¿y en otras herramientas de filología digital?

5 minutos de preparación y, si es oportuno, pueden mostrar sitios web o imágenes de proyectos (Marina, Víctor, Víctor... ☺)



José Manuel Frajedas

Primerísimos pasos...

- Tesis doctoral «Tratado de cetrería»: texto, gramática y vocabulario (según el Ms. 9 de la R.A.E.), que publica en 1997 en formato HTML en servidores de la UNED.
- Participó en la primera mesa redonda sobre Filología e Informática en el primer congreso de Historia de la Lengua.
- Transcripción en soporte informático de los tratados de cetrería españoles de la Edad Media (algunos fueron publicados en formato de microficha por el HSMS en 1992).

A continuación...

- En 2010 publica [La codificación XML/TEI de textos medievales](#) tras descubrir la TEI gracias a Charles Faulhaber.
- Emplea herramientas como *Juxta*, para la colación automática de textos, *ExistDB* para la publicación de textos etiquetados en TEI.

Nuevos retos...

- En 2016 publica [Lingüística forense y crítica textual. El caso Ayala–Cervantes](#), en donde utiliza herramientas de colación automatizada e identificadores de plagio para demostrar la correcta autoría de la edición del *Libro de la caza de las aves* de Pero López de Ayala, publicado por la Biblioteca Virtual Miguel de Cervantes.
- En 2015 comienza a aprender el lenguaje de programación R y la estilometría ([El análisis estilométrico aplicado a la literatura española las novelas policiacas e históricas](#)).
- Libro digital: [Cuentapalabras. Estilometría y análisis de texto con R para filólogos](#)

Para este curso...

- Desde 2020 emplea *Transkribus* en el proyecto 7PartidasDigital y pronto en uno nuevo que se titulará «La lengua de los cajistas», cuyo objetivo será la transcripción semiautomática de un amplio número de incunables en lengua castellana para su análisis lingüístico.



Francisco Gago Jover



Francisco Gago-Jover es catedrático de español en el College of the Holy Cross (Worcester, Massachusetts). Cursó estudios en la Universidad de Valladolid, recibiendo una licenciatura en Geografía e Historia en 1985. Posteriormente cursó estudios de doctorado en Lingüística y Filología hispanorromance en la Universidad de Wisconsin-Madison. Su tesis doctoral fue un estudio del léxico militar castellano en la Edad Media. Es autor de tres diccionarios de terminología militar, diversos artículos sobre lexicografía, creación de corpus lingüísticos y estilometría, y un importante número de transcripciones paleográficas de textos medievales españoles.

Diccionario de la prosa castellana del Rey Alfonso X
†Lloyd A. Kasten, John J. Nitti, Francisco Gago Jover
Hispanic Seminary of Medieval Studies - © 2024

introducción fuentes abreviaturas

lemas | formas

a1
a2
aabruz
aanca
aaptamiento
aauquoz
saram
sayn
ab
abacero
abad
abadesa
abadia
abajado
abajador
abajamiento
abaldonar
abanto
sbarca
abastado
abastamiento
abastar
abastecer
abatimiento
abatir
abatritaz
abece
abeja
abertura

a1
stut.
1 * Letra del abecedario.
LAP (1250?-1279?) fol. 1r33, CCN: 00000001, azcrizat dizen en griego ala, quarta piedra dela **a**.
GE1 (1272-1275) fol. 166v-88, CCN: 00000002, pues si contaremos en el a.b.c. fallaremos que es la primera letra.

2 • Usado para designar un punto geométrico.
AST (1277) fol. 167r-46, CCN: 00000003, sea la linea que fue puesta de primero; linea de **a**.b. & el punto que fue puesto en ella; el punto de **a**. & los dos puntos que fueron puestos a su diestro & a su siniestro.

Formas atestiguadas: **a**, ha, haber₂.
Véase también: **a**, ha, haber₂.

a2
prep.
1 * Denota el complemento de la acción del verbo, ya precediendo a nombres, ya a otros verbos en infinitivo.
LAP (1250?-1279?) fol. 1r33, CCN: 00000004, et assi fizieron otros muchos sabios que en estas cosas tamxieron, mas entre aquelllos ouy v algunos que se metieron mas a saber el fecho dellas.
LAP (1250?-1279?) fol. 1r34, CCN: 00000005, et en el suelo de la tierra se puso un gran mullido cada uno aquel terrazo por otro nuevo. & desque fuere quemada & la sacaren del forno. & pusieren cabo dela un poco de sufre, & rucieren la piedra con agua; saldra della fuego tan fuerte que quemara muy lueno a toda cosa que alcance que pueda quemar.

LAP (1250?-1279?) fol. 9v12, CCN: 00000006, et a otra uertud esta piedra que si la atan con filo de seda uermeja al pescuezo a mugier que sea premnada quel tanga al uiente, guarda la criatura fastal tiempo del parlimento que non reciba mal nin danno; si non fuere por fuerça.

LAP (1250?-1279?) fol. 29v26, CCN: 00000007, en tierra de egipto las fallan en las mineras que ya diximons o son falladas las otras piedras en unas cuevas que muy fondas que son y, en que crían muchas bestias fieras & poncónadas, assi que los que entran dentro a tomar las diafanias por ellas.

LEY (1265-1277) fol. 167r-46, CCN: 00000008, & puede demandar aquil a quien fue fecho el donadio que polo cumplian, si en alguna cosa fuere minraud, fuera ende si fuese plazo punto para mesurar aquella cosa quel diessen, & no quisiese uenir a tomar la **a** medir la, o si el se touliesse antes por pagado de lo que diessen, & se dexasse de la quanta.

CRZ (1259) fol. 37v44, CCN: 00000009, mas si esta coniunction sobrechida fuere en alguno de los signos que son signifidadores de los enemigos, que el rey significa que mataran aquen enemigo, & ad aquel contrallador daquel rey, & quel rey tomara ueangana del, & que folgara del mal sabor & del pesar quel fazia.

GE4 (1280) fol. 258v67, CCN: 00000010, qui ama al senyor rogar la por sos peccados, & astener sa & guardar se dellos, & en su oracion auia fazio en voz diaz eza ouida.

A lo largo de los años ha ofrecido numerosos talleres, presenciales y en línea, sobre diferentes aplicaciones de las humanidades digitales (paleografía y transcripción automática de textos, estilometría, diseño de corpus lingüísticos, etc.).

Como Director de Proyectos Digitales del Hispanic Seminary of Medieval Studies, tiene a su cargo la *Biblioteca Digital de Textos en Español Antiguo*, la bibliografía *Lexical Studies of Medieval Spanish Texts*, y el *Old Spanish Textual Archive*, un corpus lingüístico, lematizado y etiquetado morfológicamente, de cerca de 35.000.000 de palabras, de textos medievales escritos en castellano, asturiano, leonés, navarro-aragonés y aragonés.

Programa

Curso: Filología digital: transcripción automática de manuscritos, colación asistida y creación de corpus textuales

HORA*	LUNES	MARTES	MIÉRCOLES	JUEVES	VIERNES
9h00 - 10h30	Introducción al curso (Mónica Castillo Lluch)	Transcripción automática de manuscritos medievales (Mónica Castillo Lluch)	Colación asistida de diversas versiones textuales con CollateX (Mónica Castillo Lluch)	7PartidasDigital: edición de mss e impresos con Transkribus (José Manuel Frajedas Rueda)	Introducción a la creación y aprovechamiento de los corpus textuales (Francisco Gago Jover)
10h30 - 12h	Filología tradicional y digital (Mónica Castillo Lluch)	Transcripción automática de manuscritos modernos (Mónica Castillo Lluch)	Colación asistida de diversas versiones textuales con CollateX (Andrea Escobar Castillo)	7PartidasDigital: análisis estilométrico (José Manuel Frajedas Rueda)	Creación de corpus textuales (Francisco Gago Jover)
12h - 12h30	Descanso				
12h30 - 14h	Preparación técnica de los talleres del curso (Andrea Escobar Castillo)	Taller eScriptorium (Andrea Escobar Castillo)	Análisis de datos tras la colación con CollateX (Mónica Castillo Lluch)	Taller Transkribus (José Manuel Frajedas Rueda)	Aprovechamiento de los corpus textuales (Francisco Gago Jover)
14h - 15h	Comida				
15h - 16h30		Taller eScriptorium (Andrea Escobar Castillo)	Taller CollateX (Andrea Escobar Castillo)	Taller Transkribus (José Manuel Frajedas Rueda)	
16h30 - 18h.		Taller eScriptorium (Andrea Escobar Castillo)		Sesión de discusión: proyectos y problemas (Francisco Gago Jover)	

Filología tradicional y digital

Filología tradicional



REAL ACADEMIA ESPAÑOLA



ASOCIACI N DE ACADEMIAS DE LA
LENGUA ESPA OLA

Consulta posible gracias al compromiso
con la cultura de la



por palabras

Escriba aqu  la palabra

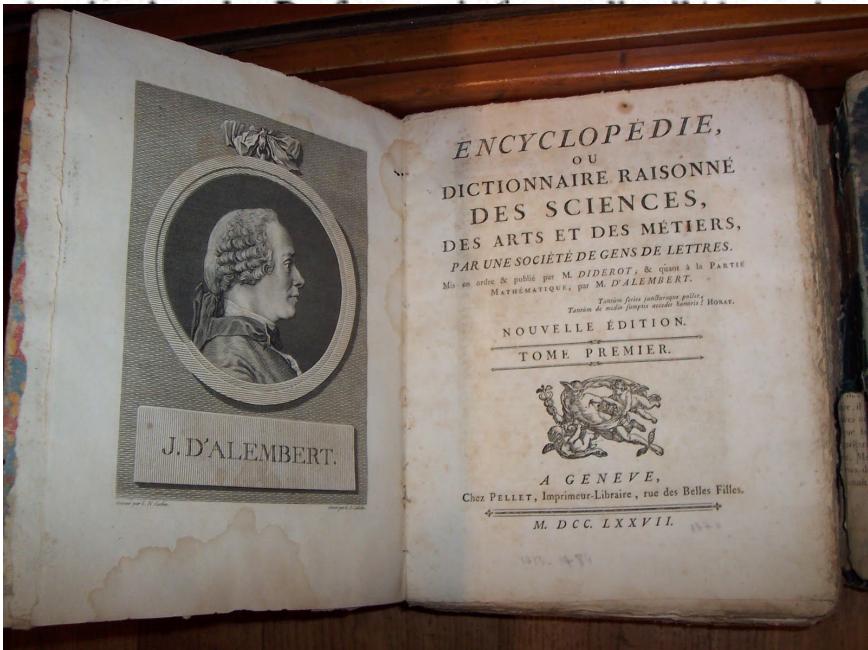
Consultar

filolog a

Del lat. *philología*, y este del gr. φιλολογία *philología*.

1. f. Ciencia que estudia las culturas tal como se manifiestan en su lengua y en su literatura, principalmente a trav s de los textos escritos.
2. f. T cnica que se aplica a los textos para reconstruirlos, fijarlos e interpretarlos.
3. f. **ling stica**.

c'est le plus long siège de ville que l'on connoisse. Nabuchodonosor assujettit vraisemblablement les *Philistins* avec les autres peuples de la Syrie, de la Phénicie, & de la Palestine. Ils tomberent ensuite sous la



texte grec porte φιλοβοκτός, *Philoboctos*. (D. J.)

PHILOCANDROS, (Géog. anc.) île de la mer *Ægée*, & l'une des Cyclades, selon Ptolomée, *L. III. c. xv.* Pline, *L. IV. c. xij.* & Etienne le géographe écrivent *Philecandros*, & la mettent parmi les îles Sporades. Hesychie écrit *Phlegandros*. On la nomme aujourd'hui *Policandro*: elle est entre les îles de Milo & de Sikino. (D. J.)

PHILONIUS PORTUS, (Géog. anc.) port de l'île de Corse. Ptolomée, *L. III. c. ij.* le place sur la côte méridionale près d'*Alistia*. Niger & Léander dis-

nant relâche à ses courriers fatigués, *Philogée*, Pyrocis, Eous & Ethon, cherche les bosquets d'Amphirite pour se reposer lui-même avec les nymphes océanides. Il baigne ses rayons à moitié plongés, & tantôt montrant un demi-cercle doré, il donne un dernier regard lumineux, & disparaît enfin totalement dans le sein de Téthis. (D. J.)

PHILOLAUS, (Mythol.) Esculape avoit un temple près de la ville d'*Alope* dans la Laconie, où il étoit honoré sous le nom de *Philolaüs*, c'est-à-dire *bon & salutaire aux hommes*. Il ne pouvoit avoir un surnom plus glorieux. (D. J.)

PHILOLOGIE, f. f. (*Littérat.*) espèce de science composée de grammaire, de poétique, d'antiquités, d'histoire, de philosophie, quelquefois même de mathématiques, de médecine, de jurisprudence, sans traiter aucune de ces matières à fond, ni séparément, mais les effleurant toutes ou en partie.

Ce mot est dérivé du grec φίλος & λόγος, amateur des discours, des lettres ou des sciences.

La *philologie* est une espèce de littérature universelle, qui traite de toutes les sciences, de leur origine, de leur progrès, des auteurs qu'elles ont cultivées, &c. Voyer POLYMATHIE.

La *philologie* n'est autre chose que ce que nous appelons en France les *Belles-lettres*, & ce qu'on nomme dans les universités les *humanités*, *humaniores litteræ*. Elle faisoit autrefois la principale & la plus belle partie de la Grammaire. Voyer GRAMMAIRE & GRAMMARIEN.

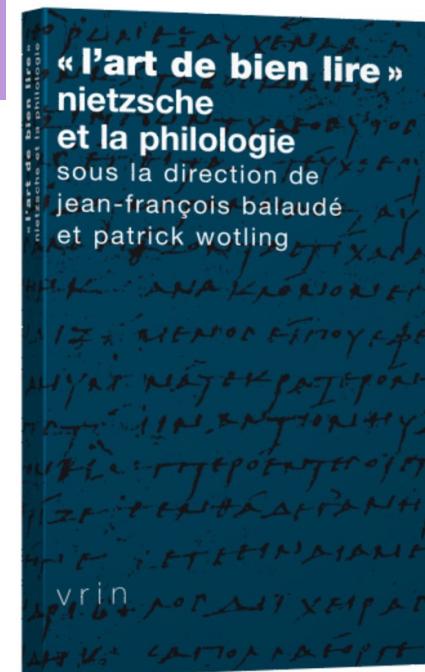
PHILOLOGUE, f. m. (*Littérat.*) on appelle ainsi

La philologie est l'art de bien lire.

L'art de pouvoir faire lire. « La tâche du philologue consiste à travailler des textes anciens pour les offrir à notre compréhension. »

Le philologue s'assure de l'identité de l'auteur d'un texte, l'authenticité de son contenu, il propose une contextualisation linguistique et historique sans laquelle un texte demeure obscur.

2012



Nietzsche, la philologie comme pratique et comme métaphore

Vendredi 18 août 2023

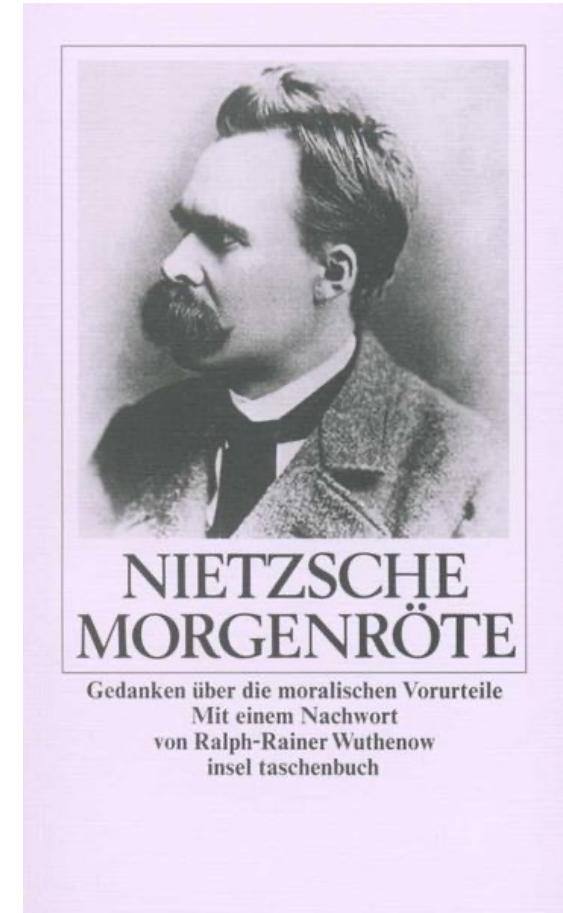
▶ REPRENDRE (6 MIN)



[Podcast France Culture](#)

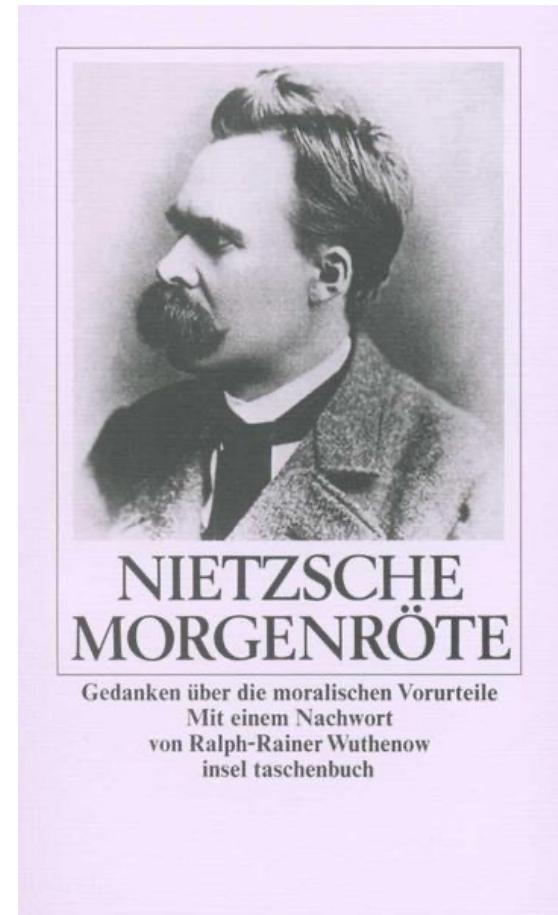
Jetzt gehört es nicht nur zu meinen Gewohnheiten, sondern auch zu meinem Geschmacke — einem boshaften Geschmacke vielleicht? — Nichts mehr zu schreiben, womit nicht jede Art Mensch, die „Eile hat“, zur Verzweiflung gebracht wird.

Philologie nämlich ist jene ehrwürdige Kunst, welche von ihrem Verehrer vor Allem Eins heischt, bei Seite gehn, sich Zeit lassen, still werden, langsam werden —, als eine Goldschmiedekunst und -kennerschaft des Wortes, die lauter feine vorsichtige Arbeit abzutun hat und Nichts erreicht, wenn sie es nicht lento erreicht. Gerade damit aber ist sie heute nötiger als je, gerade dadurch zieht sie und bezaubert sie uns am stärksten, mitten in einem Zeitalter der „Arbeit“, will sagen: der Hast, der unanständigen und schwitzenden Eilfertigkeit, das mit Allem gleich „fertig werden“ will, auch mit jedem alten und neuen Buche: — sie selbst wird nicht so leicht irgend womit fertig, sie lehrt gut lesen, das heißt langsam, tief, rück- und vorsichtig, mit Hintergedanken, mit offen gelassenen Türen, mit zarten Fingern und Augen lesen ...



Ahora no sólo forma parte de mis hábitos, sino también de mi gusto -¿un gusto travieso quizás? - No escribir nada más que no lleve a la desesperación a todo tipo de persona que tenga "prisa". La filología, al fin y al cabo, es ese arte venerable que exige sobre todo una cosa a su admirador -hacerse a un lado, tomarse su tiempo, sosegarse, volverse lento- como un arte de orfebre y conocedor de la palabra, que tiene que hacer todo un trabajo fino y cuidadoso y no consigue nada si no lo consigue lento. Pero precisamente por eso es hoy más necesaria que nunca, precisamente por eso nos atrae y encanta con más fuerza, en medio de una época de "trabajo", es decir: de prisas, de prisas indecentes y sudorosas, que quieren "acabar" todo de una vez, incluso con cada libro viejo y nuevo: -ella misma no acaba nada tan fácilmente, enseña a leer bien, es decir a leer despacio, profundamente, hacia atrás y con cuidado, con segundas intenciones, con las puertas abiertas, con los dedos y los ojos tiernos...

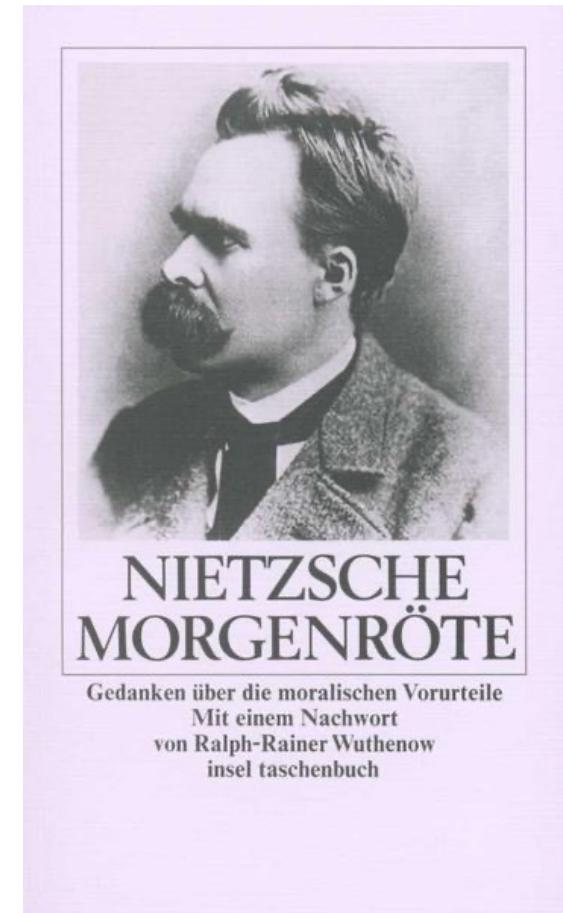
[Translated with DeepL](#)



1886

Ahora no sólo forma parte de mis hábitos, sino también de mi gusto -¿un gusto travieso quizás? - No escribir nada más que no lleve a la desesperación a todo tipo de persona que tenga "prisa". La filología, al fin y al cabo, **es ese arte** venerable que exige sobre todo una cosa a su admirador -hacerse a un lado, tomarse su tiempo, sosegarse, volverse lento- como un arte de orfebre y conocedor de la palabra, que tiene que hacer todo un trabajo fino y cuidadoso y no consigue nada si no lo consigue lento. Pero precisamente por eso es hoy más necesaria que nunca, precisamente por eso nos atrae y encanta con más fuerza, en medio de una época de "trabajo", es decir: de prisas, de prisas indecentes y sudorosas, que quieren "acabar" todo de una vez, incluso con cada libro viejo y nuevo: -ella misma no acaba nada tan fácilmente, **enseña a leer bien**, es decir a leer despacio, profundamente, hacia atrás y con cuidado, con segundas intenciones, con las puertas abiertas, con los dedos y los ojos tiernos...

[Translated with DeepL](#)

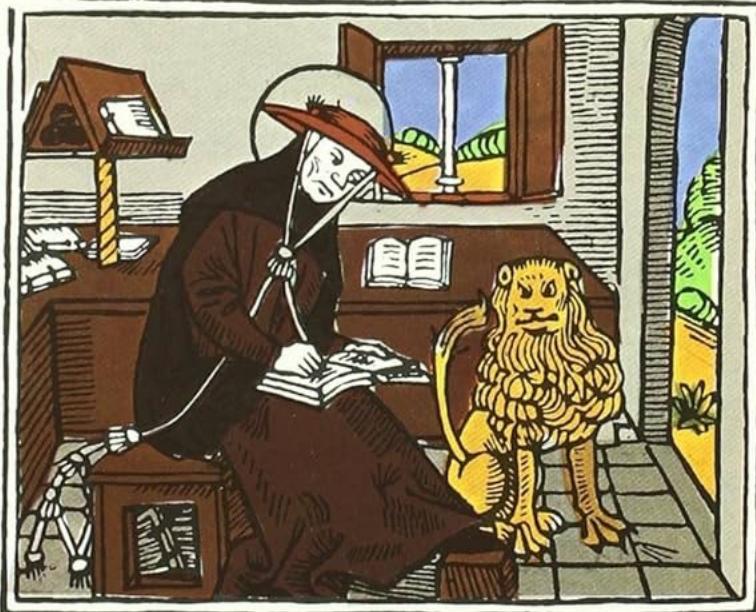


1886
arte

3. m. o f. Conjunto de preceptos y reglas necesarios para hacer algo.

ALBERTO BLECUA

Manual de crítica textual



EDITORIAL CASTALIA

El objetivo de la filología consiste en reconstruir la palabra original.

La búsqueda de un método que permitiera eliminar en lo posible lo subjetivo se remonta a los orígenes de la filología y se halla estrechamente ligada a la concepción del libro como transmisor de unos significados fundamentales —la *res*—, que se conformaban a su vez a través de unos significantes únicos —los *verba*—. Cualquier alteración de estos últimos podía provocar catástrofes exegéticas irremediables, en particular en aquellos textos que transmitían, nada menos, la palabra divina. La Biblia, la *Iliada*, Aristóteles o Platón, pilares de una cultura basada en la palabra escrita y en su interpretación, debían permanecer incorruptibles, inmutables, siempre idénticos a su estado original. Una mala interpretación bíblica por error de una letra podía crear un hereje; un error en el nombre de una planta en Dioscórides podía ocasionar una defunción. Si la filología bíblica reconstruía la palabra divina, la filología clásica reconstruía la palabra pagana, comenzando, naturalmente, por la Gramática. En este último caso, un error mínimo —por ejemplo, en un régimen verbal— podía alterar la reconstrucción del sistema. El Brocense, con gran escándalo de cierto sector de

(Blecua 1983: 9)



Moisés de Miguel Ángel (1513-1515)
Basílica de San Pietro in Vincoli

¿Por qué tiene dos cuernos?

Explicación de [National Geographic](#)

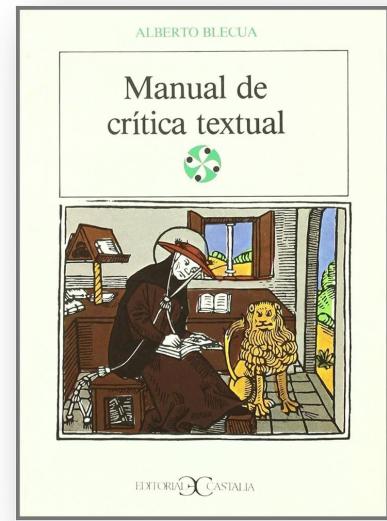
A finales del siglo IV San Jerónimo de Estridón, padre y doctor de la Iglesia, hizo un monumental trabajo traduciendo la Biblia desde el griego y el hebreo al latín.

Cuando San Jerónimo tradujo Éxodo 34, 35 se encontró con una palabra compuesta por las letras **KRN**. (Es importante aclarar que en hebreo no se escriben las vocales.) La cual se puede interpretar como «*keren*» (radiante, luminoso, con rayos de luz) o «*karan*» (cuerno).

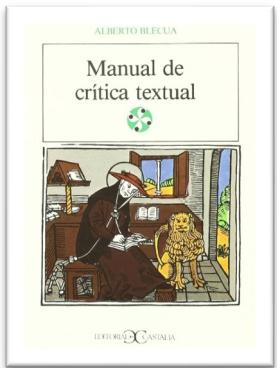
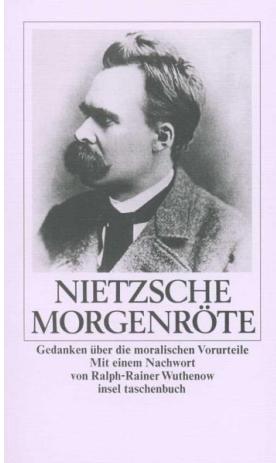
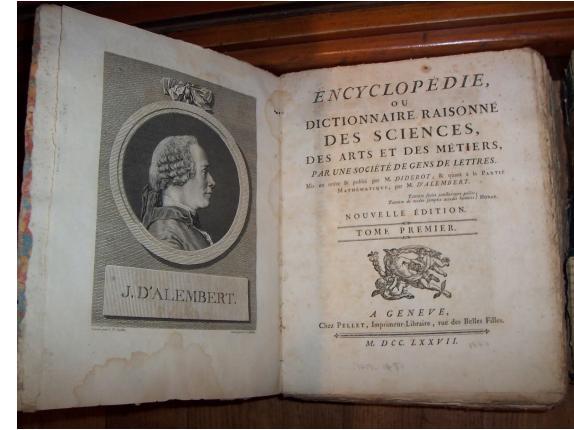
- Actualmente la mayoría de las Biblias traducen aquel pasaje de la siguiente manera:
“Y los Hijos de Israel vieron entonces que rayos de Luz emanaban de la tez del rostro de Moisés”.
- San Jerónimo, en cambio, optó por la siguiente traducción:
- *“Y los Hijos de Israel vieron entonces que cuernos emanaban de la tez del rostro de Moisés”.*

Desde fechas lejanas se fueron aplicando los mismos métodos de la filología bíblica o clásica a los textos vulgares. Su fin: reconstruir el sistema lingüístico y cultural de cada 'nación' desde sus orígenes, y mantener la palabra original de un 'autor', que, por otra parte, es lo que se había intentado hacer siempre con los clásicos para su correcta interpretación. Pero la filología vul-

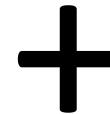
(Blecua 1983: 10)



philologie, s. f. (*Littérat.*) : espèce de science composée de grammaire, de poétique, d'antiquités, d'histoire, de philosophie, quelquefois même de mathématiques, de médecine, de jurisprudence, sans traiter aucune de ces matières à fond ni séparément, mais les effleurant toutes ou en partie.
(Diderot et d'Alembert, 1775, *Encyclopédie*)



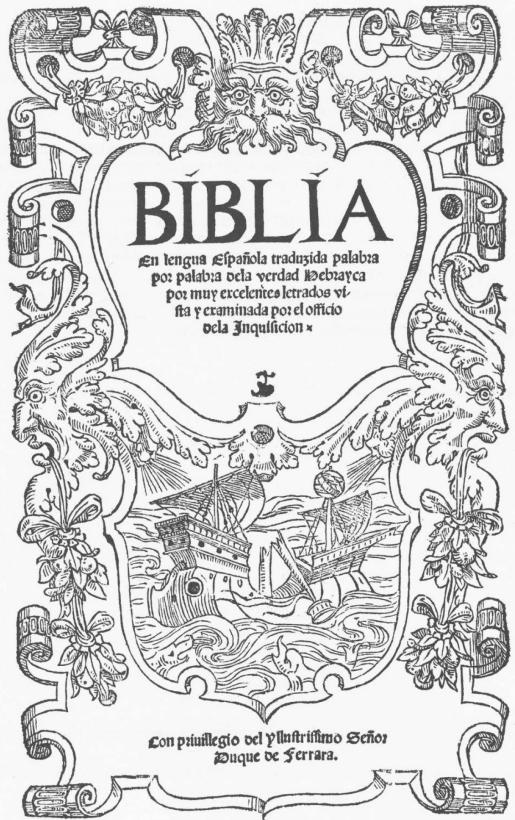
Reconstruir la palabra original



El arte que enseña a leer bien

Podríamos definir la filología como

- el arte que enseña a leer bien los textos,
- con ayuda de conocimientos de disciplinas diversas para interpretarlos,
- para permitir a los demás que lean esos textos.



Biblia de Ferrara

Traducción en judeoespañol de la Biblia hebrea (el Tanaj) por Abraham Usque y Yom Tom Atias, publicada en Ferrara en 1553, para uso de los sefardíes.



MS 206

Hebrew square book script. Iraq. 1st half of 11th c.

Tanaj, Wikipedia

la qual no se dero por otra por no quitar a cada uno lo suyo. Y no piense ninguno que la lectura della es como la de los otros libros que de vna o dos veces se entienden: por que segun dijen los sabios que toda lection se deve leer diez vezes primero q se diga que se leyó : que se leyó dice y no que se entienda quanto mas la sagrada escriptura

«...dizen los sabios que toda lectión se deve leer diez veces primero que se diga que se leyó; que 'se leyó' se dice, y no que se entienda.»

(*Biblia de Ferrara*, en el prólogo)

(Blecua 1983: 9)

ADVERTENCIA PRELIMINAR

«...dizen los sabios que toda lección se deve leer diez veces primero que se diga que se leyó; que 'se leyó' se dice, y no que se entienda.»

(*Biblia de Ferrara*, en el prólogo)

La crítica textual es un arte que ofrece una serie de consejos generales extraídos de una práctica plurisecular sobre los casos individuales de naturaleza muy diversa. Por ejemplo, analizando múltiples casos particulares de los errores que cometan los copistas en diferentes épocas, lenguas y géneros, es posible reducir el error de copia a una tipología limitada. Parece claro que quien conozca el arte de editar textos estará, en el peor de los casos, en mejores condiciones de llevar a cabo una edición que aquel que ignore los presupuestos mínimos.

La búsqueda de un método que permitiera eliminar en lo posible lo subjetivo se remonta a los orígenes de la filología y se halla estrechamente ligada a la concepción del libro como transmisor de unos significados fundamentales —la *res*—, que se conformaban a su vez a través de unos significantes únicos —los *verba*—. Cualquier alteración de estos últimos podía provocar catástrofes exegéticas irremediables, en particular en aquellos textos que transmitían, nada menos, la palabra divina. La Biblia, la *Iliada*, Aristóteles o Platón, pilares de una cultura basada en la palabra escrita y en su interpretación, debían permanecer incorruptibles, inmutables, siempre idénticos a su estado original. Una mala interpretación bíblica por error de una letra podía crear un hereje; un error en el nombre de una planta en Dioscórides podía ocasionar una defunción. Si la filología bíblica reconstruía la palabra divina, la filología clásica reconstruía la palabra pagana, comenzando, naturalmente, por la Gramática. En este último caso, un error mínimo —por ejemplo, en un régimen verbal— podía alterar la reconstrucción del sistema. El Brocense, con gran escándalo de cierto sector de

que segun dizen los sabios que toda lección se deve leer diez veces primero q se diga que se leyó : que se leyó dize y no que se entienda quanto mas la sagrada escriptura

¿Cambios en la cita/edición de Blecua?

(Blecua 1983: 9)

ADVERTENCIA PRELIMINAR

«...dizen los sabios que toda lección se deve leer diez veces primero que se diga que se leyó; que ‘se leyó’ se dice, y no que se entienda.»

(*Biblia de Ferrara*, en el prólogo)

La crítica textual es un arte que ofrece una serie de consejos generales extraídos de una práctica plurisecular sobre los casos individuales de naturaleza muy diversa. Por ejemplo, analizando múltiples casos particulares de los errores que cometan los copistas en diferentes épocas, lenguas y géneros, es posible reducir el error de copia a una tipología limitada. Parece claro que quien conozca el arte de editar textos estará, en el peor de los casos, en mejores condiciones de llevar a cabo una edición que aquel que ignore los presupuestos mínimos.

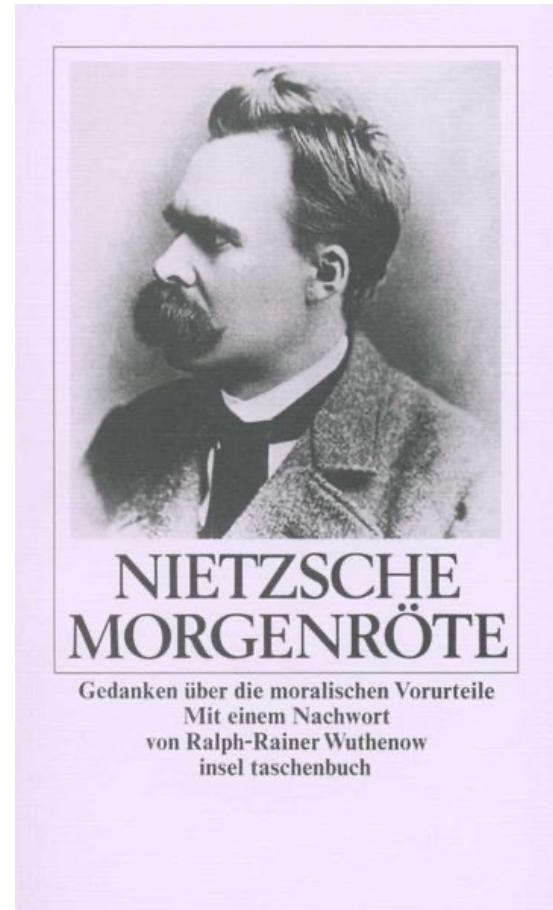
La búsqueda de un método que permitiera eliminar en lo posible lo subjetivo se remonta a los orígenes de la filología y se halla estrechamente ligada a la concepción del libro como transmisor de unos significados fundamentales —la *res*—, que se conformaban a su vez a través de unos significantes únicos —los *verba*—. Cualquier alteración de estos últimos podía provocar catástrofes exegéticas irremediables, en particular en aquellos textos que transmitían, nada menos, la palabra divina. La Biblia, la *Iliada*, Aristóteles o Platón, pilares de una cultura basada en la palabra escrita y en su interpretación, debían permanecer incorruptibles, inmutables, siempre idénticos a su estado original. Una mala interpretación bíblica por error de una letra podía crear un hereje; un error en el nombre de una planta en Dioscórides podía ocasionar una defunción. Si la filología bíblica reconstruía la palabra divina, la filología clásica reconstruía la palabra pagana, comenzando, naturalmente, por la Gramática. En este último caso, un error mínimo —por ejemplo, en un régimen verbal— podía alterar la reconstrucción del sistema. El Brocense, con gran escándalo de cierto sector de

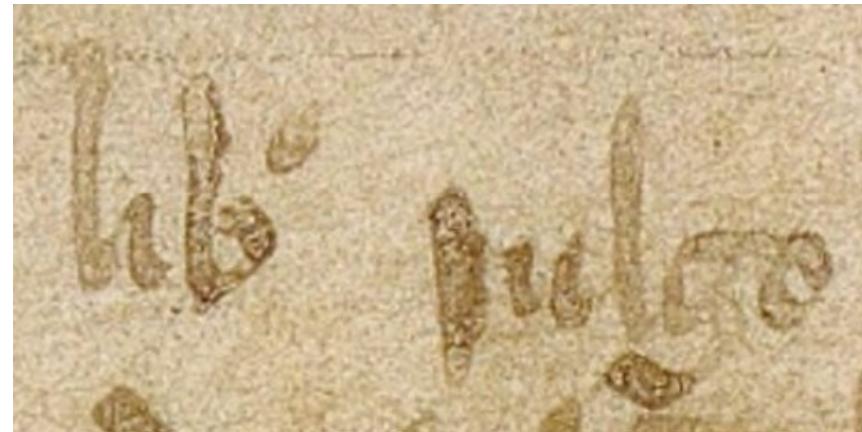
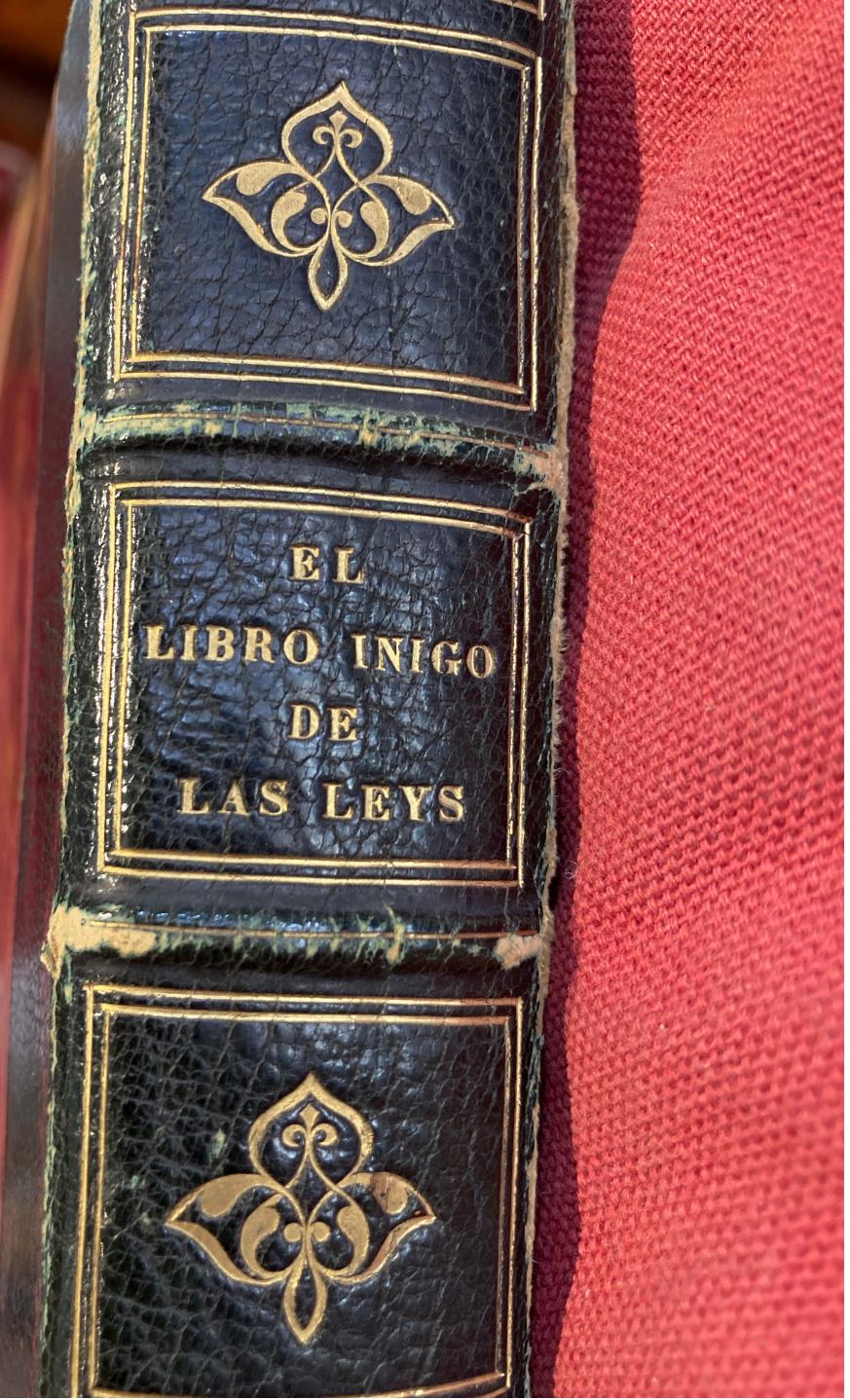
que segun dizen los sabios que toda lección se deve leer diez veces primero q se diga que se leyó : que se leyó dize y no que se entienda quanto mas la sagrada escriptura

¿Cambios en la cita/edición de Blecua?

- acentuación
- puntuación → para facilitar la lectura
- comillas
- *dize > se dize*
- no se moderniza la grafía original (salvo s largas)

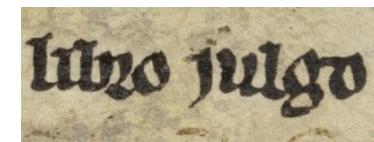
Philologen: lernt mich gut lesen!
(1886, en [introducción a Morgenröte](#))





BNF esp. 256

« Libro julgo » (Fuero juzgo)
(en algunos manuscritos occidentales)

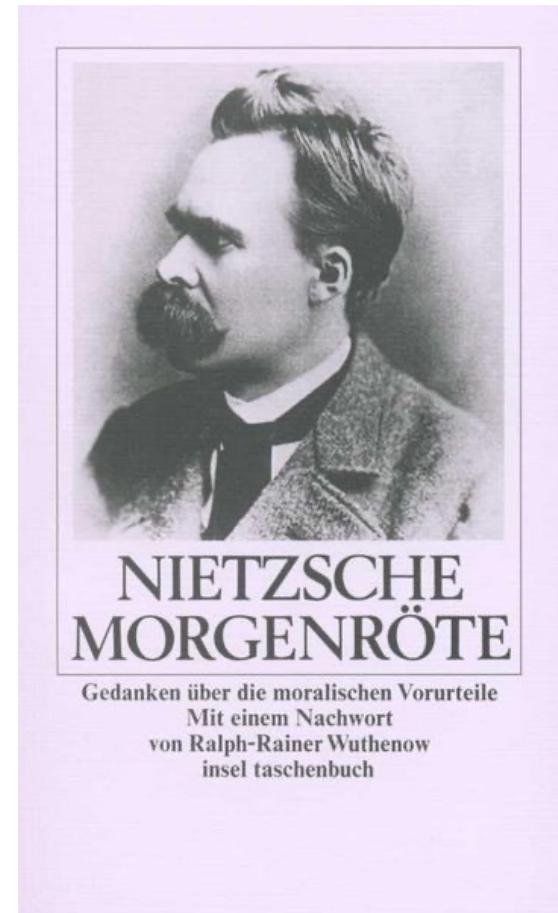


BNE5814

BNF esp. 256 (error en el título del lomo de la encuadernación por mala lectura)

Ahora no sólo forma parte de mis hábitos, sino también de mi gusto-¿un gusto travieso quizás?- No escribir nada más que no lleve a la desesperación a todo tipo de persona que tenga "prisa". La filología, al fin y al cabo, es ese **arte** venerable que exige sobre todo una cosa a su admirador-hacerse a un lado, tomarse su tiempo, sosegarse, volverse **lento-** como un arte de orfebre y conocedor de la palabra, que tiene que hacer todo un trabajo fino y cuidadoso y no consigue nada si no lo consigue **lento**. Pero precisamente por eso es hoy más necesaria que nunca, precisamente por eso nos atrae y encanta con más fuerza, en medio de una época de "trabajo", es decir: de **prisas**, de prisas **indecentes y sudorosas**, que quieren "**acabar**" todo de **una vez, incluso con cada libro viejo y nuevo**:-ella misma no acaba nada tan fácilmente, enseña a leer bien, es decir a leer **despacio, profundamente**, hacia atrás y con cuidado, con segundas intenciones, con las puertas abiertas, con los dedos y los ojos tiernos...

[Translated with DeepL](#)



Leemos a Nietzsche convencidos de que tiene razón en recomendar la lectura lenta...

[Manifiesto de Johannes Kabatek](#)

Slow linguistics – a manifesto

Johannes Kabatek, University of Zurich

kabatek@rom.uzh.ch

The following lines should not be misunderstood. What I want to propose here is not an extension of the tents of the ‘slow food’ movement to the academy in general and to linguistics in particular¹. It is, rather, a continuation of what I argued in a paper in Spanish on “empathic linguistics”² and has to do with the special status of the study of human language within the scientific disciplines, although some of the main thoughts are valid for work in all of them. The issues briefly raised here are not in themselves new, but they are of course controversial. I offer them as a starting point for discussion, and I am looking forward to receiving comments!

... pero, somos individuos de nuestro tiempo (vivimos en una era aún más acelerada que la de Nietzsche en la que las máquinas permiten procesos con una rapidez inimaginable hace 150 años) y queremos:

- emprender proyectos para terminarlos
- utilizar las herramientas de filología digital que ayudan al filólogo

Compromiso del filólogo hoy: mantener la exigencia de la buena lectura con la ayuda de la filología digital, poniendo a prueba de modo crítico esas herramientas y con lectura lenta y rigurosa para garantizar la calidad de los textos transcritos y editados.

La filología emplea las herramientas a su alcance (entre ellas las digitales, como cualquier disciplina hoy).

Operar con estas herramientas requiere la misma prudencia y control que en medicina, biología, finanzas...

El filólogo es el mediador entre los textos y la sociedad

- lee bien
- edita textos a partir de testimonios
- contribuye a la creación y mejora de corpus de textos
- estudia los textos desde el punto de vista lingüístico y cultural



RAE54, 3v: escriba trabajando en su mesa

Filología digital

“the collation of manuscripts requires the infuriating accuracy of a pedant and the obsessive stamina of an idiot. It is therefore an ideal task for a computer” (Robinson 1989: 99).

Robinson, Peter M. W. (1989): «The Collation and Textual Criticism of Icelandic Manuscripts (1): Collation», *Literary and Linguistic Computing* 4.2, 99-105.



¿Qué plusvalía le ves tú a la filología con herramientas digitales?

Ventajas:

productividad

+

unificación del patrimonio

accesibilidad

difusión...

El filólogo es el mediador entre los textos y la sociedad

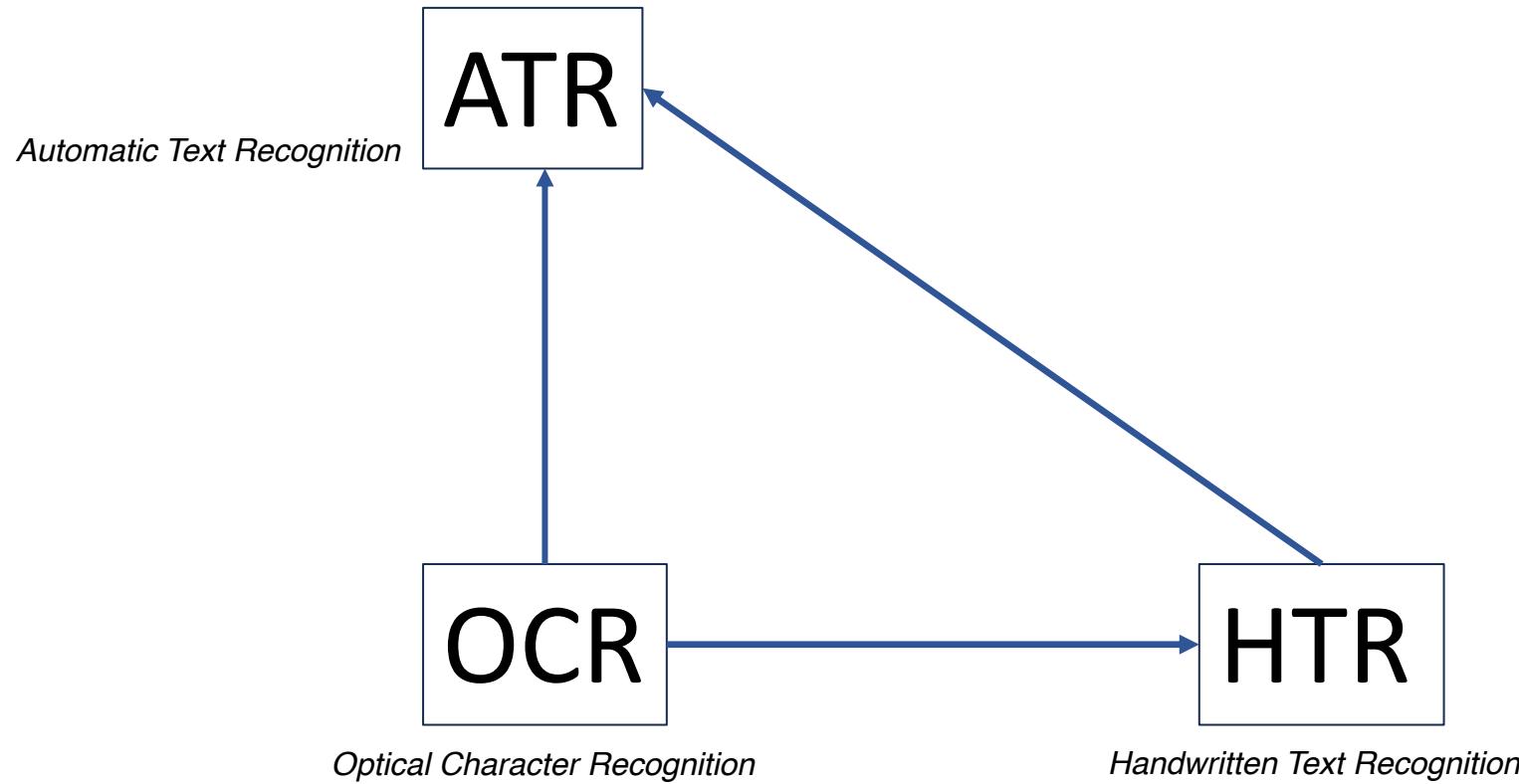
- lee bien
- edita textos a partir de testimonios
- contribuye a la creación y mejora de corpus de textos
- estudia los textos desde el punto de vista lingüístico y cultural

- Lectura con programas de transcripción automática
- Colación asistida para establecer un texto crítico
- Edición con herramientas digitales (TEI)
- Mediante criterios de edición comunes, referenciación de metadatos
- Etiquetado gramatical, léxico, semántico, gráfico... (datos procesables estadísticamente con R u otros)

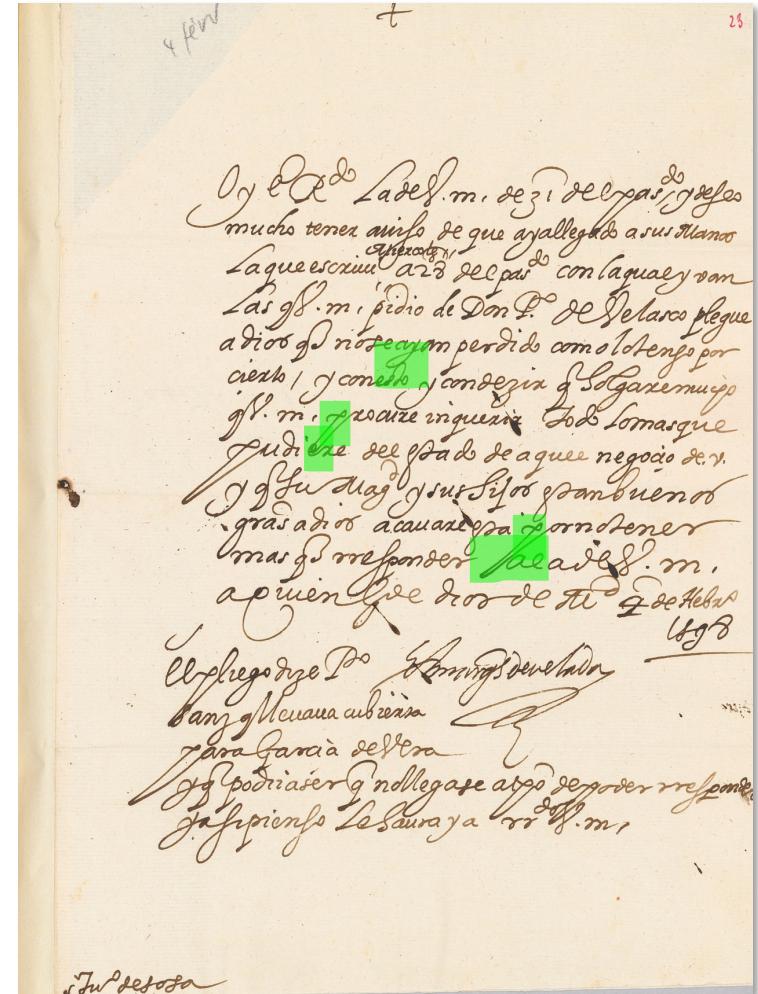


RAE54, 3v: escriba trabajando en su mesa

¿Qué?



- variabilidad de la disposición de la escritura en la página
- degradación de los documentos
- amplia variabilidad gráfica
 - escritura cursiva
- superposición de los caracteres
- sangrado y traspaso de tinta



¿Por qué?

- ahorro de tiempo y costo
- conservación de documentos históricos, facilitación de su acceso
- revela y unifica el patrimonio previamente olvidado
- contribuir a la creación de datos, base de la creación de modelos de reconocimiento

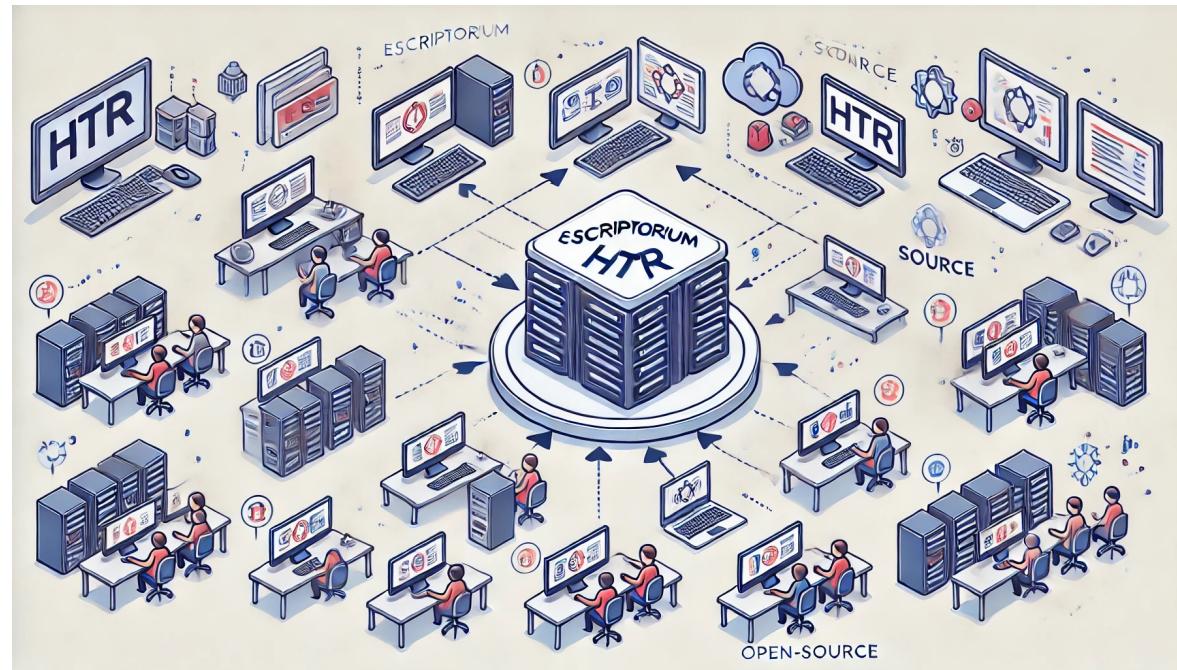


señor
mote huimino e d
memoria que en el
ya la Religion de
Semana y cuya que
acabado de su a
y Recat. los su a
Cebillo y Scottie
que se deuen esperar
bu aquisea Rech
Ma prima ya la
medis sumeladas
en el Coto del Concio
a que yo que
char los Deudos
Onera con
Alverdiacini de
el Precio de l
de Francia y di
dus Juan. Acierto
sabeler que
do. dos Cartas de V. m. de 20 y El
medize del electo de osma, Conesta
ngregaciones queseanTenido esta
era ynbiar los decretos Pero nosean
den y Perfiéonar El Latin dellos
Con mucho cuy dado y desseo que E
Como Convienec y destas Congrega
El atrauajado y trauaja harto enlo
sea ya Verlo Acabado yo le edadohai
s. y la Vltima Vez quele Able enello
para la Semana queViene
Toledo hasta El cap^o. 8 de la Action
nas Se Vera Con breuedad El espe
esera lo quemas dilacion a de Causal
ongregaciones seatratado dela
epulchro quese hallo enlay Glesia
[que] en t^po del obispo don Ber^do
ser delos martires de aquella ciu
s demas queV. m. Vera por dos
esta vno de la planta de lay gl^a
la forma y ta

Herramientas

- eScriptorium [B. Kiessling , Tissot , Stokes et Ezra 2019 y Benjamin Kiessling 2019] → motor OCR: Kraken
- Transkribus [Kahle , Colutto , Hackl y Mühlberger 2017]
 - ↳ motor OCR: Pylaia
- Calfa [Vidal-Gorène et al. 2021]

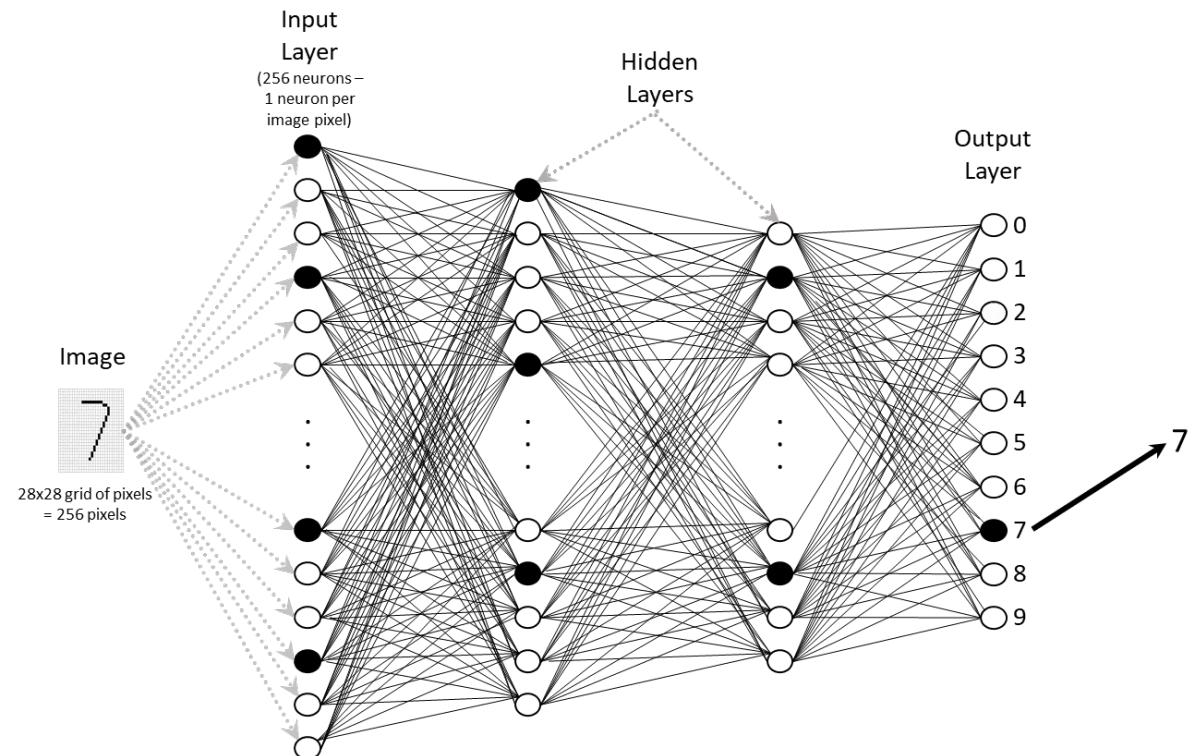
Motores OCR más usados: Tesseract, Textract (Amazon), Google Cloud Vision.



¿cómo funcionan?



Corpus: datos correctamente etiquetados a mano (gold corpus)



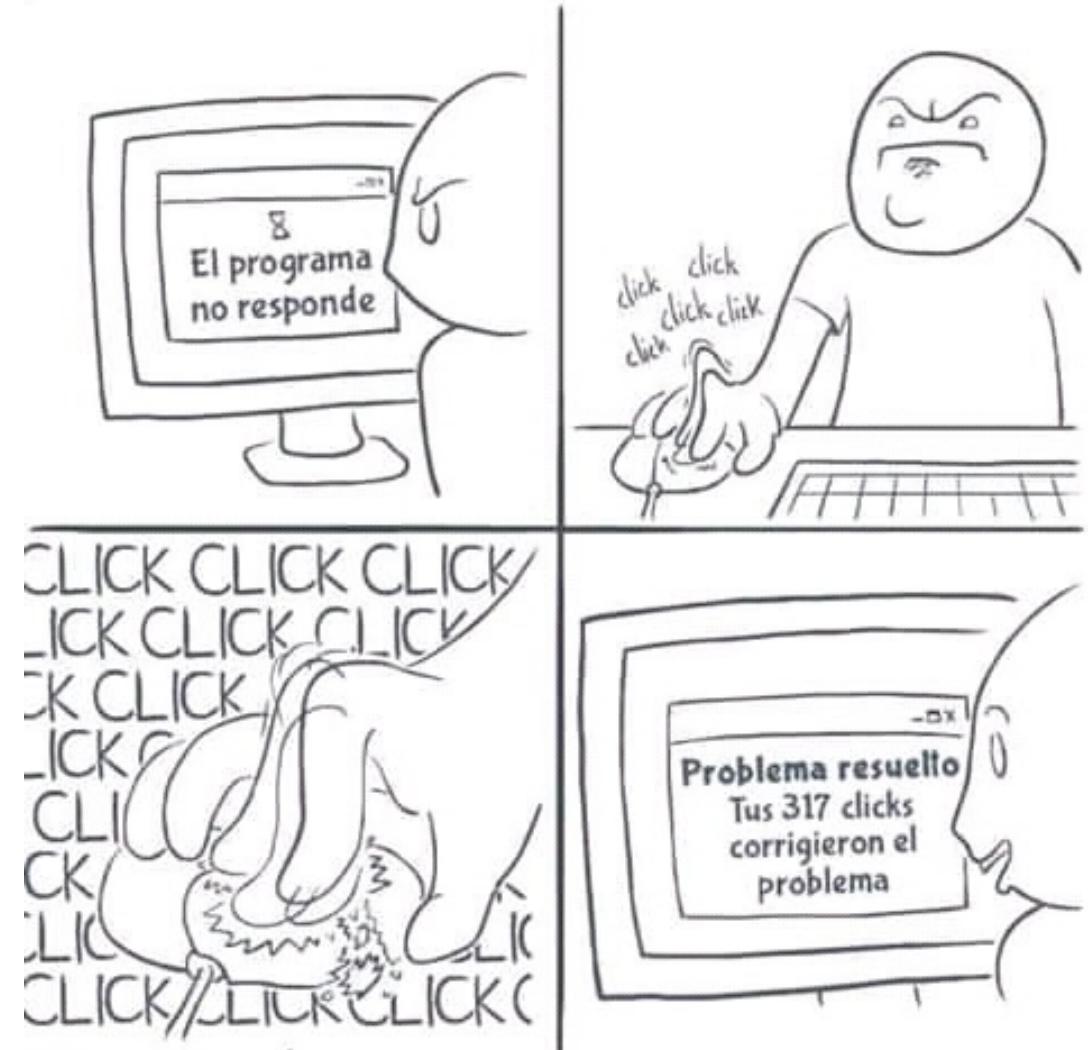
Creación de modelo a partir de redes neuronales: predicción

Una aplicación multitarea y la buena voluntad del filólogo...

- Adquisición del facsímil de mejor calidad (\neq buena resolución)
- Pre-procesamiento de las imágenes, si es pertinente
- Análisis de la página (segmentación): **identificación de zonas** (texto, ilustraciones, capitales, etc.),
identificación de líneas de texto.
- Reconocimiento: predicción de un carácter con ayuda de un modelo
- Corrección manual
 - Exportación y explotación de los datos
 - Corrección y afinamiento de modelo para mejorar el desempeño de un modelo de base

Preparación técnica de los talleres del curso

<https://tinyurl.com/yck7cdra>



Bibliografía

Camps, Jean-Baptiste (2016), *La Chanson d'Otinel : édition complète du corpus manuscrit et prolégomènes à l'édition critique*, Tesis doct, Université Paris-Sorbonne.

Gille Levenson, Matthias (2023), «Towards a general open dataset and model for late medieval Castilian text recognition (HTR/OCR)», en Pinche, Arianne y Stokes, Peter (eds.), *Journal of Data Mining and Digital Humanities. Special Issue : Historical documents and automatic text recognition*.

Kahle, Philip, Sebastian Colutto, Günter Hackl y Günter Mühlberger (2017), «Transkribus-a Service Platform for Transcription, Recognition and Retrieval of Historical Documents». En *14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, IEEE;, vol. 4, 19-24.

Kiessling, Benjamin (2019), «Kraken - an Universal Text Recognizer for the Humanities». DH2019 : Complexity. 9-12 July 2019, Utrecht, The Netherlands.

Kiessling, Benjamin, Robin Tissot, Peter Stokes y Daniel Stökl Ben Ezra (2019), «eScriptorium: An Open Source Platform for Historical Document Analysis». En *2019 International Conference on Document Analysis and Recognition Workshops (ICDARW)*, vol. 2.

Vidal-Gorène, Chahan, Noémie Lucas, Clément Salah, Aliénor Decours-Perez y Boris Dupin (2021), «RASAM. A Dataset for the Recognition and Analysis of Scripts in Arabic Maghrebi». En *International Conference on Document Analysis and Recognition*, Springer, 265-281.

Rodríguez Díaz, Elena E. (2024), «Elementos para fechar los códices leoneses y castellanos según los manuscritos datados (ss. XII y XIII)», en Ángeles Romero Cambrón (ed.): *La ley de los godos. Estudios selectos*. Bern, Peter Lang.