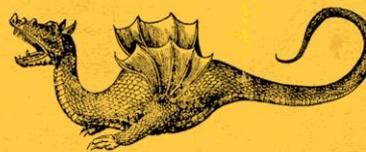




PROGETTO
MAMBRINO

HISTORIAS FINGIDAS



De editor analógico a editor digital

José Manuel Fradejas Rueda

(Universidad de Valladolid)*

Abstract

En el artículo se cuentan los recientes desarrollos del proyecto colaborativo 7PartidasDigital. En concreto, se señalan los principales rasgos del trabajo filológico tradicional y se apuntan los posibles avances que pueden ofrecer los medios informáticos a la hora de realizar una edición crítica digital. Entre los aspectos más destacados, insistimos en la transcripción automatizada de testimonios impresos y en el uso de sistemas de colación automática, dos recursos que de manera concreta agilizan y simplifican el trabajo filológico.

Palabras clave: Humanidades Digitales; *Siete Partidas*; Alfonso X el Sabio; filología digital

The present article is centered on the recent developments of the collaborative project 7PartidasDigital. Specifically, the main aspects of traditional philological work and the possible advances that computational tools can offer in the making of a digital critical edition are pointed out. Among the most outstanding aspects, we insist on the automated transcription of printed testimonies and the use of automatic collation systems, two resources that concretely speed up and simplify philological work.

Keywords: Digital Humanities; *Siete Partidas*; Alfonso X the Sage; digital philology



* Este trabajo forma parte de los resultados del proyecto *7PartidasDigital* (referencias FFI-2016-75014 y PID2020-112621GB-I00) cuyo objetivo es una edición crítica digital de las *Siete Partidas*. Este proyecto <<https://7partidas.hypotheses.org/>> (cons. 03/05/2022) se desarrolla desde la Universidad de Valladolid y cuenta con la financiación de la Agencia Estatal de Investigación del Reino de España y se integra dentro de la Red de Excelencia ‘Cultura escrita medieval hispánica: del manuscrito al soporte digital (CEMH)’ (RED2018-102330-T), Agencia Estatal de Investigación - Ministerio de Ciencia e Innovación.

En los comienzos del proyecto 7PartidasDigital, la Biblioteca Nacional de España me invitó a hablar en una jornada sobre la edición digital¹. En esa ocasión verbalicé un sueño: poder tomar las imágenes digitales de un texto que hubiera en un repositorio y que gracias a un flujo de trabajo adecuado con las herramientas informáticas pertinentes pudiera obtener al final una edición crítica digital.

En 2021 estamos un poco más cerca de cumplir ese sueño, porque durante el último año hemos añadido al equipamiento filológico digital dos nuevas herramientas que facilitan el objeto final de la investigación: ofrecer la edición crítica digital de las *Siete Partidas*, aunque aún estamos lejos de llegar al final. Las herramientas incorporadas son la transcripción automatizada de testimonios impresos y un sistema de colación automática.

Antes de exponer cómo trabajan los editores digitales, quiero exponer cómo trabajaba como editor analógico. No es que un día me planteara dejar el mundo analógico y entregarme al digital. *Primum philologia, deinde computatrum scientia*; no se puede ser editor digital sin una sólida base de filología y crítica textual.

En 1979, me inicié en la crítica textual con un texto de principios del siglo XIV muy breve, apenas 28 folios a dos columnas, y con un total de casi 21000 palabras-token²: el ms. 9 de la Real Academia Española que contiene un *Tratado de cetrería*; un manuscrito encuadernado junto con una copia del XVIII que hizo Antonio de Santiago y Palomares, y del que localicé otra copia dieciochesca en la Biblioteca de la Fundación Universitaria Española (Archivo Campomanes 52-2) que había realizado el mismo Palomares. Por último, hallé el antecedente de la copia del trescientos, un grueso manuscrito de la Real Biblioteca de San Lorenzo de El Escorial, el ms. V-II-19 y el antecedente de una parte de este: el hoy RES/270 de la BNE, por aquel entonces en manos de unos anticuarios que estaban vendiendo algunos de los códices de la colección de Sir Thomas Phillipps (ms. 11719). El ms. de la RAE no era una copia del

¹ No existe publicación, pero en la red se puede visionar la comunicación <<https://www.youtube.com/watch?v=QWlfSqtvGt4&t=1100s>> (cons. 28/10/2021).

² Se entiende por palabra-token cualquier secuencia de caracteres alfanuméricos entre dos espacios en blanco o signos de puntuación; es decir, las *palabras ortográficas*.

escurialense, sino que, a partir de las seis obras recopiladas por una misma mano y con una misma letra, se había creado el texto antológico que se conserva en el manuscrito académico.

Por aquel entonces la filología española, a pesar de los grandes éxitos de la escuela pidaliana, no había producido un corpus teórico de referencia sobre crítica textual en el que los neófitos pudieran aprender algo que no se enseñaba en las aulas ni en los seminarios universitarios. Tan solo se contaba con un artículo publicado en 1917 por Américo Castro³ y, casi tres lustros más tarde, otro artículo de Marín Ocete (1932), en el que exponía el estado actual de la crítica de textos. Para comprender la metodología empleada por la escuela de filología española había que recurrir a las introducciones de algunas ediciones preparadas por algunos de los más egregios miembros del Centro de Estudios Históricos como las que el mismo Castro menciona en su artículo, o la de Onís (1912) y, especialmente, la de García Solalinde (1930). Aunque de estos trabajos se infiere que lo que les preocupaba a sus autores era establecer «textos fidedignos» o «ediciones fidedignas» (Santiago, 2020, 228, n.15) o «ediciones dignas» (Santiago, 2020, 231), a pesar de que, como indica Fernández-Ordoñez (1988, 231), «don Ramón [...] creía en el carácter científico y objetivo del método» de la crítica textual.

Así, pues, el único medio que había para aprender cómo hacer una edición de un texto (no me atrevo a decir que crítica) era leer y examinar las ediciones de textos medievales. De esas lecturas inferí un procedimiento que, con el tiempo y la aparición en 1983 del *Manual de crítica textual* de Alberto Blecuá, fui perfeccionado y enriqueciendo desde el punto de vista teórico. Este manual llegó tarde para mí, pues defendí la tesis (Fradejas Rueda, 1983) el mismo año en el que el *Manual* salió de las prensas⁴.

Años más tarde se estableció que en el *Diccionario de términos filológicos* de Fernando Lázaro Carreter (Santiago, 2005) se recogía la primera

³ Reeditado siete años después (Castro, 1924).

⁴ Wehrli (1951), López Estrada (1979) y Jauralde Pou (1981) dedicaron un breve capítulo informativo en sus manuales de introducción a los estudios literarios, como lo haría, años más tarde Ruiz (1985), esta a la vista ya del manual de Blecuá (1983). Había mucho más movimiento en este sentido en la Universidad de Buenos Aires, en el Seminario de Crítica Textual fundado por Germán Orduna y en el que se comenzó en 1981 a editar la revista *Incipit*.

descripción del neolachmanianismo. Pero un diccionario no es el mejor lugar para aprender técnicas y métodos de investigación porque lo primero, quien lo hubiera pretendido, tendría que haber conocido de antemano cuáles eran esos términos técnicos. Además, la primera edición (Lázaro Carreter, 1953) solo incluía seis: *aparato crítico*, *colacionar*, *edición*, *fijación de un texto*, *filología* y *variante*. En la segunda (Lázaro Carreter, 1961) aumentó la nómina de términos hasta la cuarentena (Santiago, 2005). Tampoco se contaba con formación ni fuentes de información codicológica. El magnífico *Manual de codicología* de Elisa Ruiz no apareció hasta 1988.

La lectura y estudio del manual de Blecua (1983) confirmó que el barrunto metodológico no había sido muy equivocado. La verdad es que, a grandes rasgos los principios básicos para la edición crítica de un texto romance –no puedo hablar de otros ámbitos lingüísticos– son los mismos y son prácticamente inamovibles.

Lo primero es establecer el catálogo de las copias de la obra que se pretende editar. Por entonces, la *Bibliografía de la literatura hispánica* de Simón Díaz (1950-1993)⁵ era la obra de referencia básica donde se iniciaban las búsquedas, aunque, como veremos, para el ámbito castellano medieval, en 1975, aparece una bibliografía que cambiará totalmente el panorama (Cárdenas *et al.*, 1975).

El siguiente paso era obtener las reproducciones de los testimonios. Hasta la primera década del siglo XXI lo que se obtenían eran micropelículas, en blanco y negro y en formato de 35 mm, aunque a veces podían ser de 16 mm. La mayor dificultad era que solían ser películas positivas, no negativas, por lo que no se podía ir a un laboratorio fotográfico del barrio y encargar que las ampliaran sobre papel fotográfico. Había que utilizar unas voluminosas máquinas que permitían leerlas. No obstante, poco a poco fueron apareciendo las que podían, con tan solo pulsar un botón, reproducirlas sobre papel, como si de una fotocopia se tratara. El último gran avance fue cuando se pudieron escanear los microfilms, obtener imágenes JPEG y manejarlas posteriormente con un ordenador.

⁵ Los volúmenes dedicados a la Edad Media se publicaron en 1963 y 1965. En 1972 se inició la publicación de los tomos dedicados a los Siglos de Oro

Ya fueran ampliadas e impresas en papel, ya fueran leídas en la pantalla, había que transcribirlas a mano. Es decir, se realizaba una nueva copia del manuscrito y sobre esta copia, con muchísimo cuidado, se plasmaban todas las anotaciones necesarias. Un excelente ejemplo de este sistema son los cuadernos en los que Rafael Lapesa esbozó en los años 1950 su edición, jamás publicada, del *Rimado de Palacio* de López de Ayala (Lapesa, 2010).

Una vez hecha la transcripción y la labor crítica, había que pasarlo a máquina, mecanografiarlo. Pero las máquinas de escribir tenían unas posibilidades gráficas muy limitadas. Las máquinas de escribir españolas, por lo general, carecían de <ç> y no digamos de los imprescindibles corchetes y, por supuesto, de la ese alta <ſ> y la nota tironiana (el *ampersand* <&>, un excelente sustituto, no estaba en teclado español usual), por lo que había que dibujarlos a mano (la <ç> se obtenía al escribir una <c>, retroceder el carro una posición y sobrescribir una coma)⁶.

Es decir, seguíamos inmersos en un mundo de transmisión manuscrita, sujeta a los mismos errores mecánicos de copia que sufrieron los copistas que escribieron los códigos objeto de nuestro estudio. Pero no hacíamos una única copia, sino dos, al menos, la manuscrita y la mecanografiada, con lo que la tipología y las posibilidades de error de copia se ampliaban.

En agosto de 1981, IBM lanzó al mercado el IBM Personal Computer. En este momento es cuando realmente se inició la popularización del ordenador. Es cierto que ya existía el Apple II (1977) y otras máquinas como los Sinclair ZX Spectrum⁷ o el Commodore Vic-20 (1980). Estos dos últimos, casi juguetes, fueron la plataforma de lanzamiento, pero los IBM y los clónicos que le siguieron fueron los que lograron que los ordenadores fueran accesibles a los especialistas de humanidades.

Al principio, no pasaron de ser potentísimas máquinas de escribir y con funciones de almacenaje. Los procesadores de texto, ya fueran Wordstar, WordPerfect o Word, cambiaron el modo de editar los textos,

⁶ Tampoco tenía los números 1 y 0 para los que se usaba la <l> minúscula y la <O>.

⁷ El Sinclair ZX Spectrum fue una evolución del Sinclair ZX80, lanzado en 1980, y del Sinclair ZX81, de 1981.

puesto que había un sistema de almacenaje y modificación de lo copiado que permitía una fácil corrección, sin tener que volver a copiar todo el folio. Los bloques de texto podían moverse de un lado a otro de los documentos, o guardarlos en ficheros separados para reutilizarlos cuando hiciera falta. No todo era perfecto, al principio el juego de caracteres fue muy limitado, pero pronto se amplió para incorporar la mayoría de las letras con diacríticos de la mayoría de las lenguas europeas occidentales.

Téngase en cuenta que estoy hablando de las pocas posibilidades de un profesor de universidad sin fondos para investigación y sin acceso a los centros de computación. Solo contaba con una grandísima curiosidad. Sin embargo, en otras latitudes están empleando los ordenadores para la edición de textos y otras fruslerías filológicas (Nitti, 1978). Basta leer los capítulos de Hockey (1980, 149-167) acerca de los procedimientos de edición de textos a principios de los años 1980 para darse cuenta de por dónde iba el mundo.

Aunque ya había acceso a los ordenadores, como he dicho, los percibíamos potentes máquinas de escribir y la forma de trabajar en crítica textual se ha mantenido, salvo contadas excepciones, casi invariable hasta la segunda década del 2000. Quizá mejorara un poco debido a las muchas facilidades que los ordenadores e internet fueron ofreciendo con el paso del tiempo.

A la hora de editar la versión castellana medieval de la *Epitome rei militaris* (Fradejas Rueda, 2011) seguí empleando los viejos sistemas de colación⁸. Una vez decidido cuál era el texto que iba a usar de base, procedí a su transcripción y corrección en pantalla de todos los errores que se hubieran deslizado en este proceso. A continuación, lo imprimí en hojas A3 y, con un juego de rotuladores ultrafinos, con una amplia paleta de colores, un bloc de notas autoadhesivas para apuntar las notas que surgían a lo largo del proceso de colación y una caja de hojas de etiquetas adhesivas –de las que se suelen usar para las etiquetas de dirección postal– para corregir los errores que pudiera cometer al anotar variantes, procedí a la *collatio* de los testimonios conservados. Este mismo procedimiento lo

⁸ Varios autores han descrito los sistemas de colación manual que utilizaron (Manley y Rickert, 1940; West, 1973) e incluso algunos de si eran sistemas sólidos (Moorman, 1975, 47) o no (Foulet y Speer, 1979, 48).

utilicé poco después para la edición del *Libro de la caza de las aves* de Pero López de Ayala (Fradejas Lebrero y Fradejas Rueda, 2016). Este caso fue un poco más complejo porque se trataba de obra con una cuarentena de testimonios.

Sin embargo, desde mis orígenes como editor filológico, me interesé por los métodos informáticos. Al principio, se limitaban a un sistema de transcripción semipaleográfica codificada, que el Hispanic Seminary of Medieval Studies (HSMS) creó para el que quizá sea el proyecto pionero de las humanidades digitales hispánicas, aunque ideado por hispanomedievalistas norteamericanos de la Universidad de Wisconsin - Madison (Buelow y Mackenzie, 1986): el *Dictionary of Old Spanish Language* (DOSL), del que una de las derivadas más conocidas y utilizadas es la Bibliografía Española de Textos Antiguos (BETA) dentro del base de datos bibliográfica *Philobiblon*, un catálogo de fuentes primarias para la literatura medieval peninsular⁹.

Este sistema lo apliqué al texto del ms. 9 de la RAE y a todos los textos cetreros castellanos que fui localizando y transcribiendo a lo largo del tiempo. Además, desarrollé pequeños programas en BASIC para hacer las concordancias y los cálculos de frecuencia.

Años más tarde, los textos se publicaron dentro del corpus textual de Madison, en un sistema esotérico e inimaginable hoy día: en microfichas que requerían localizar una máquina que pudiera leerlas. Con el tiempo, se publicaron en CD-ROM, como sucedió con el corpus del *scriptorium alfonsí* (Kasten, Nitti, Jonxis-Henkemans, 1997) y el más amplio de manuscritos e impresos antiguos del HSMS (O'Neill, 1999).

La lección que extraje de esta época inicial, algo que siempre ha sido mi mayor preocupación, fue la absoluta necesidad de reutilizar los materiales y su perdurabilidad. Esto no siempre ha sido fácil porque, hasta

⁹ La primera edición de este catálogo estuvo a cargo de Cárdena, Nitti y Mackenzie (1975). La última versión impresa, la tercera, fue cosa del equipo encabezado por Faulhaber *et al.* (1984). Se diseñaron las normas para una cuarta edición (Faulhaber y Gómez Moreno, 1986) pero se abandonó el proyecto analógico para convertirse en otro informático. Al principio como parte del proyecto *Admyte* (1991), del que se independizaría y adquiriría vida como un gran catálogo en línea que integra tres bibliografías: gallego-portuguesa (BITAGAP), catalana (BITECA) y española (BETA) bajo el nombre general de *Philobiblon* (Faulhaber, 1997).

la estandarización del sistema MS-DOS, cada fabricante utilizaba su propio sistema operativo. El sistema operativo CP/M (Control Program for Microcomputers) era casi específico para cada fabricante o grupo de fabricantes. Bajo este sistema operativo se desarrolló el procesador de textos Wordstar y la base de datos dBase, por lo que pasar de una máquina a otra podía ser problemático.

En mi caso, empecé con un ordenador Toshiba T-100 con CP/M, y cuando quise cambiar de ordenador, a un IBM PC, tuve que buscar una pasarela que me permitiera pasar los datos desde discos grabados bajo CP/M al sistema de archivos de IBM-DOS. El traspaso de los datos fue a través de dos máquinas diferentes: del Toshiba T-100 a un Osborne 1, pues este tenía ordenador una rutina que permitía leer los discos CP/M del Toshiba y, una vez en formato Osborne, mediante un Decision Mate V de NCR, que podía leer el CP/M de Osborne, los pude transformar en MS-DOS (IBM-DOS) sin pérdida de datos. Desde ahí han llegado a ordenadores Apple y desde discos de 5.25” de 160 Kb, a discos de 3.5” de 1.4 Mb, a memorias USB y, finalmente a *drives* en la nube. Es un problema que hay que tener en cuenta. Hoy casi lo tenemos solucionado, pero muchas cosas de gran interés como Admyte (1992-1998) o la magnífica edición del *Poema del Cid* de la BNE (Manuscrito 1998) han quedado obsoletas puesto que no se actualizaron.

Personalmente, a lo largo de estos años de cambios, he participado en varios proyectos de informatización, digitalización y edición de textos medievales españoles. Algunas de las transcripciones que preparé para el HSMS (Fradejas Rueda, 1992a; 1992b) se incorporaron a Admyte (1992-1998). Participé en la Colección Clásicos Tavera (Fradejas Rueda, 1999), que consistía en la digitalización de los objetos físicos (manuscritos e impresos) a base de convertir los microfilm de la BNE en imágenes jpeg e incluirlas en CD-ROM. Por su parte, el Centro Virtual Cervantes creó una colección de ediciones digitales en línea y me encargaron la de la *Historia de Enrique fi de Oliva* (Fradejas Rueda, 1997). En todos estos casos, actué como editor filológico; no estaba entre mis cometidos la parte tecnológica, pero no quiere decir que no estuviera al tanto de las posibilidades.

En la segunda mitad de los años 90 se abrió un nuevo mundo:

cualquiera podía crear páginas web. En la UNED, que era la universidad en la que era profesor en aquellos momentos, comencé a desarrollar, más como *hobby* que como proyecto de investigación, una web en la que editar y publicar todos los textos de cetrería que había ido transcribiendo y editando a lo largo de los años.

Cuando en el año 2000 me trasladé a la Universidad de Valladolid, reubiqué en sus servidores de la Facultad de Filosofía y Letras aquella página web. Sorprendió bastante al personal del centro de cálculo (STIC) que uno «de letras» quisiera hacer una página web para un proyecto. Al principio hubo algo de reticencia, pero lo conseguí¹⁰.

A raíz de los trabajos para la declaración de la cetrería como Patrimonio Intangible de la Humanidad diseñé un proyecto que presenté durante el congreso *Falconry: A World Heritage*, celebrado en Abu Dhabi en septiembre de 2005, para el que solicité financiación al MEC / MINECO al año siguiente (referencias HUM2006-0932/FILO y FFI20210-15128). Lo titulé *Archivo Iberoamericano de Cetrería* (www.aic.uva.es). Hoy sigue vivo, aunque en fase letárgica porque no hay materiales que añadir.

¹⁰ No quedan rastros de esas viejas versiones, salvo lo que se puede localizar en la Wayback Machine, un sistema que almacena instantáneas de sitios web y hoy se pueden recuperar con cierta facilidad, como en este caso <<https://web.archive.org/web/20010520162153/http://gramola.fyl.uva.es/~cetreria/>> (cons. 29/10/21), que ofrece el acceso a cómo estaba la página el 20 de mayo de 2001.



ARCHIVO IBEROAMERICANO DE CETRERÍA
ARQUIVO IBEROAMERICANO DE CETRARIA

introducción - catálogo - legal



| | | |
|---|--|--|
| <p>CATALOGO</p> <p>MANUSCRITOS</p> <p>IMPRESOS ANTIGUOS</p> <p>IMPRESOS MODERNOS</p> <p>REVISTAS</p> <p>ARTÍCULOS</p> <p>IMÁGENES</p> <p>DOCUMENTACIÓN</p> |  | <p>Libro de cetrería que hyzo Euangelista corriendo fortuna por el golfo de Leon, a Dios msericordia, por no estar vçioso. Y trata de las aues de rrapyna, de los talles y plumajes y propiedades de cada vna, y de los goujernos y curas para sus dolencias como adelante oytes.</p> <p style="text-align: center;">Capitulo prymero. Del falcon gyryfalte.</p> <p>Ffalcon gyryfalte se dize por falcon que jura falso, y tales son ellos Causalos que se engendran dentro en el huevo y naçen dentro en el nido y asy se cryan jasta que saben bolar. Tu, caçador, que lo jas de comprar, mjra que sea de talle de lanterna y el plumaje de color de su madre; y que tenga cabeça con su pico y que la tenga en cabo del pescueço, porque la sy la toujese en medio paresçeria que tenga dos papos y cabeça no njinguna. Et sobre todo mjra que sea abiuado porque los muertos apruevan muy mal en Castilla, que salen floxos desmayan y pierden el comer jasta que se secan. Son aves muy bien acondicionadas que mejor las contentarejls con vn papo de gallina que con otro vazyo de vaca. Salen dellos muy buenos golondryneros que como las golondrynas son de muy grande esfuerço sy se pusiere en deffensa firiere al tu falcon enel loga[r] vergonçoso, toma la dicha golondryna y mascala pongela en la llaga, sy con esto no sanare querria mas vn marauedy.</p> <p style="text-align: center;">Capitulo de los sacres</p> <p>El falcon sacre es asy llamado porque el primero que caço con ellos fue vn sacrystan. Sy oujeres de comprar alguno, mjra que sea muy apañado, de fechura de gujsopo. su plumaje de vna colorçilla qual mas te agrade y que tenga dos pies, en cada pie quatro dedos, y en cada dedo su vña, et sy no</p> |
|---|--|--|

© 2005, Archivo Iberoamericano de Cetrería - Arquivo Iberoamericano de Cetraria
 Instituto de Estudios de Iberoamérica - Universidad de Valladolid
 Asociación Española de Cetrería y Conservación de Aves de Rapina
 Creado 06.09.05. Última actualización 06.09.05

Fig. 1. Edición del primer folio del *Libro de cetrería* de Evangelista según el Mss/21549 de la BNE en la primera versión del *Archivo Iberoamericano de Cetrería* (2005)

En el diseño original, aspiraba a crear una edición que se asemejara a lo que se había desarrollado en el disco 1 de Admyte (1992-1998): editar los textos junto con las imágenes. Esto se desarrolló en HTML puro, un lenguaje presentacional en el que los datos y los aspectos visuales van unidos. Además, se diseñó con un sistema de marcos, por lo que el mantenimiento era complicado y la adición de nuevos materiales un difícil (Fig. 1).

Esta tarea se simplificó con el uso de hojas de estilos en cascada (CSS), sistema que separa el contenido (los textos) del aspecto visual. Esta fue otra lección muy importante en mi evolución como editor digital: mantener separados el contenido y el aspecto visual. A esto último se le suele dar mucha importancia hoy en día; algo que no ha preocupado mucho a un buen número de proyectos de edición digital, los cuales nacen

con una visión de lo que quieren obtener como producto final y no tanto de las fases intermedias, que son las básicas.

Charles Faulhaber, entonces director de la Bancroft Library (Universidad de California, Berkeley) en respuesta a uno de mis correos, dirigió mi atención hacia un sitio web de la Universidad de Berkeley: *Digital Scriptorium*¹¹. Ahí fue donde descubrí la Text Encoding Initiative y sus posibilidades¹². Esto fue especialmente revelador porque uno de los ejemplos con los que los desarrolladores ilustraban su metodología era un texto medieval castellano, con lo que pude ver las posibilidades que ese sistema tenía. *Digital Scriptorium* y los motores de búsqueda me llevaron al proyecto Medieval Nordic Text Archive¹³ (MENOTA) y sus tres capas de transcripción: 1) facsimilar, 2) paleográfica y 3) normalizada. Este proyecto me llevó, a su vez, al Medieval Unicode Font Initiative (MUFI)¹⁴, algo imprescindible para algunas de las posibilidades que abordaba MENOTA.

La TEI era la respuesta a los problemas con los que me había topado hasta entonces en la edición de textos medievales e inicié la redacción, tras un curso del TEI@Oxford Summer Schools en 2010, de una especie de manual con ejemplos castellanos de todas aquellas etiquetas que creí que eran interesantes para mi idea de una edición digital de un texto medieval castellano (Fradejas Rueda, 2010).

Por esos años, daba por agotado mis potenciales contribuciones sobre el vastísimo campo de los libros de cetrería españoles y comencé a diseñar un nuevo proyecto que rompiera textualmente con el anterior. La chispa iniciadora de este cambio de rumbo fue la presentación de *Alma littera* (Herrero de la Fuente, 2014). Uno de los artículos de este volumen

¹¹ <<https://digital-scriptorium.org/>> (cons. 28/10/21). Los contenidos actuales poco tienen que ver con lo que había durante la primera década del siglo XXI. Se puede explorar por medio de Wayback Machine: <<https://web.archive.org/web/20070609225805/>; https://www1.columbia.edu/sec/cu/libraries/bts/digital_scriptorium/technical/ds-xml/transcription_dtd/documentation/toc.html> (cons. 28/10/21).

¹² Faulhaber ya había hablado sobre la TEI en España, pero sus ideas quedaron soterradas en las actas del congreso de la lengua española de Sevilla de 1992 (Faulhaber, 1994).

¹³ <https://www.menota.org/EN_forside.xhtml> (cons. 28/10/21). Este proyecto se inició en 2001 y en él se estableció que «Det var enighet om at arkivets primære målgruppe er forskere og studenter innenfor middelalderfilologi», pero que no descartaban que en el futuro fuera «for et bredere publikum». <https://www.menota.org/DOK_RaadsReferat2001-09-10.xml> (cons. 29/10/21).

¹⁴ <<https://mufi.info/m.php?p=mufi>> (cons. 28/10/21).

era la edición y transcripción de un fragmento de los *Bocados de Oro* (Ruiz Albi, 2014) que se encontraba en la sección de pergaminos del Archivo de la Real Chancillería de Valladolid. Allí descubrí cuatro fragmentos de las *Siete Partidas* (Fradejas Rueda, 2015), aunque después supe de dos más.

De ahí surgió el proyecto de una edición crítica digital de las *Siete Partidas*, un proyecto considerado descabellado por algunos colegas porque había cientos de testimonios (véase más adelante). Es obvio que se trataba de una tarea compleja, sin embargo, mi interés no disminuyó porque es la única gran obra de la edad media castellana que aún carece de una edición crítica; el último intento es de hace dos siglos: la edición de la Real Academia de la Historia.

Estimé que la mejor manera de conseguir la edición crítica digital de las *Siete Partidas* era por medio de la transcripción de todos los testimonios conservados, por lo que mi idea inicial, actual y posiblemente final sea una edición sinóptica digital integral alineada al estilo de la que ofrece Enrique-Arias en el proyecto *Biblia Medieval*¹⁵.

El gran problema residía en cómo, algo que había sido escrito sobre pergamino y papel entre los reinados de Alfonso X y el de los Reyes Católicos y que conoció una primera etapa trepidante y no muy bien aclarada aún, y una segunda como obra en letra de molde, entre 1491 y 1555, y que tenía letras que eran relativamente sencillas de leer para un humano medianamente entrenado, podía ser convertido en texto comprensible y manejable por los ordenadores en un tiempo razonable.

Decidido a que se tratara de una edición sinóptica integral alineada (paralela) digital, la siguiente decisión fue el uso del sistema de codificación TEI. No había otra posibilidad. Para tener un arranque abarcable y relativamente sencillo, se comenzó por los impresos. De ahí que la primera fase del proyecto (2016-2020) se centrara en las llamadas ediciones históricas, la *princeps* incunable de 1491, la renacentista de 1555 y la académica de 1807. Se iniciaron los trabajos por la edición de 1555, el *textus receptus*, la *editio vulgata*, por la sencilla razón de que se trata de la versión con validez legal y que los especialistas del derecho utilizan.

No podemos olvidar que las *Siete Partidas* es un texto jurídico que

¹⁵ <<https://www.bibliamedieval.es/>> (cons. 28/10/21).

sigue teniendo validez en los tribunales de España y América, incluido, y en especial, Estados Unidos de América, aunque desde el punto de vista filológico tuviéramos que huir de ella porque es una edición contaminada, no por una mera *consultatio* o *contaminatio* sino porque es un caso flagrante de *conflatio*¹⁶. Pero poseía un elemento de sencillez adicional: la tipografía era redonda y no gótica, como en el caso de los incunables.

Se comenzó por la edición de Gregorio López de 1555 para ahorrar el esfuerzo extra que supone el aprendizaje del sistema de codificación TEI (Fradejas Rueda, 2021b). Interesaba que los transcritores se preocuparan por los problemas del texto, no por las posibles complejidades de las etiquetas TEI. Para ello se diseñaron unas protomarcas muy elementales, lo que permitía a cada colaborador utilizar el procesador de textos con el que se sintiera más cómodo. La única restricción era seguir al pie de la letra las instrucciones.

Cada fichero solo debe contener un título; se transcribe respetando la longitud de línea; se desarrollan y marcan las abreviaturas con cursivas; las rúbricas con negritas; se unen y separan las palabras con criterio actual; se marca comienzo de folio y de columna entre corchetes; los títulos corrientes con una llave de apertura y las *littera nobilior* con una llave seguida de las letras IN y un dígito que declara la altura de la inicial expresada en número de líneas. En la eventualidad de que algo no se entendiera, o hubiera erratas tipográficas, se añade a continuación de esa palabra [sic] porque no se corrige en ningún caso lo que dice el *original*.

Estos ficheros posteriormente se convertían en otros etiquetados con TEI muy básico por medio del filtro TEIOOP5 para Open / Libre Office. Después ese fichero se procesaba con un *script* en el lenguaje de

¹⁶ La edición de López (1555) toma como base un ejemplar la de Díaz Montalvo (1491, 1501, 1528, 1542 y 1550) y la corrige con ayuda de varios manuscritos que nos son totalmente desconocidos. Se ha podido demostrar que López usa como texto base la edición de 1550 (Fradejas Rueda, en prensa), la cual deriva de la edición de 1528 (por el momento no se puede establecer si la de 1550 es copia directa de la de 1528 o de la de 1542, lo más probable). La edición de 1528 es una edición corregida a la vista de otros manuscritos, corrección que llevó a cabo Francisco de Velasco. En el caso del texto de la *Primera Partida*, tuvo como modelo para la corrección un ejemplar de la llamada redacción primitiva (Craddock, 1974) y esto se demuestra por que presenta veinticinco títulos debido al desdoblamiento del título 1.19 y la incorporación del correspondiente proemio para el nuevo título 1.20. Por lo tanto, de Velasco tuvo acceso a otros códices de otras *Partidas*, pero no puede determinarse cuáles (Fradejas Rueda, en prensa).

programación R que se ocupaba de establecer las particularidades del modelo de transcripción diseñado para el proyecto e incluía la generación del `xml:id` que individualiza cada una de las casi 2700 leyes que constituyen las *Partidas*.

En el caso del incunable de 1491 se partió de la transcripción del HSMS, se analizaron las etiquetas y marcas utilizadas y por medio de otro *script* en R se obtuvo la codificación TEI que había establecido para el proyecto 7PartidasDigital.

En el caso de la edición de 1807 se partió de la versión OCR que ofrece la Biblioteca Digital Hispánica de la Biblioteca Nacional de España. Esta se ha corregido (aún estamos ello) y codificado en TEI. Este testimonio ha dejado de tener interés en nuestro proyecto porque se trata de un *descriptus* de un grupo de códices de la BNE y no supone nada en la transmisión de las *Siete Partidas*, aunque puede tener implicaciones serias en los estudios lingüísticos que se basan en esta edición porque se trata, como en el caso de la edición de Montalvo (1491) y de López (1555), de un caso de *conflatio* en el que solo ocasionalmente se advierte el cambio de modelo o la procedencia de la lección seleccionada.

Mientras se llevaba a cabo este proceso de transcripción y codificación semiautomático, se desarrollaron las protomarcas e instrucciones para la transcripción de los testimonios manuscritos. Se han probado con MN0, MN6 y ZAB y funcionan, aunque siempre es necesario afinarlas, pues no hay manuscrito que no depare alguna sorpresa. Gracias a esto, en la nueva andadura de 7PartidasDigital (2021-2024, ref. PID2020-112621GB-I00) se aplicará el mismo procedimiento para la transcripción y codificación de todos los testimonios manuscritos de la *Primera* y *Cuarta Partida*.

El siguiente problema con el que nos enfrentamos es el de la colación de los testimonios. El corpus de la 7PartidasDigital lo constituyen unos 100 testimonios con desigual número de copias para cada una de las *Partidas* (Fradejas Rueda, 2021d). Además, un problema básico de las *Siete Partidas* es que no se conoce ningún testimonio manuscrito completo. Esto se consiguió en 1491 con la edición de Díaz de Montalvo, revisada a la vista de varios manuscritos, que desconocemos, en 1528 por Francisco Velasco y cuya edición fue tomada como testimonio de base por Gregorio

López, que la corrigió a la luz de una serie ignota de manuscritos.

| Partida | Copias | Palabras | Palabras totales |
|---------|--------|----------|------------------|
| 1 | 15 - 1 | 143.000 | 2.000.000 |
| 2 | 25 - 1 | 142.000 | 3.500.000 |
| 3 | 17 - 3 | 184.000 | 2.000.000 |
| 4 | 12 - 3 | 62.000 | 560.000 |
| 5 | 13 - 3 | 85.000 | 850.000 |
| 6 | 15 - 1 | 64.000 | 900.000 |
| 7 | 17 - 2 | 83.000 | 1.240.000 |
| Totales | 100 | 763.000 | 11.000.000 |

Tabla 1. Cálculo estimado de palabras gráficas (tokens) en los testimonios de las *Siete Partidas*

Estos cien testimonios suponen unos once millones de palabras, como puede comprobarse en la tabla 1, por lo que colacionar manualmente, como se ha expuesto con anterioridad, sería imposible.

A principios de la década de 1960, Vinton A. Dearing (1962, 20) describió un programa creado para un ordenador *mainframe*, un IBM 7090, que le permitía comparar hasta noventa y nueve textos y aventuraba que «reasonably flexible machines will in time appear on the market at prices that will make them attractive to universities [...]. It will then be possible to do textual comparisons in minutes instead of months».

La evolución de los ordenadores, tanto en potencia de procesamiento como en sencillez de manejo y abaratamiento, ha sido enorme desde entonces. Así, Hockey (1980) ofreció una reseña de las técnicas de computación utilizadas en la crítica textual en la primera época, revisión que reelaboró posteriormente y consideró que «of the many collation system that have been written, two stand out because of their sophistication and functionality»: TUSTEP y Collate (Hockey, 2000, 126). Ambos programas fueron diseñados para ser ejecutados en ordenadores personales, TUSTEP - Tübingen System von Textverarbeitungs-Programmen (Castrillo Benito, 1992) sobre MS-DOS (Windows) y Collate en los Apple Macintosh (Robinson, 1989a y 1989b). Así mismo, ambos

programas siguen existiendo y se han actualizado constantemente; sin embargo, no son programas sencillos de manejar puesto que están diseñados para llevar a cabo todas las fases necesarias para una edición crítica de un texto, incluso la preparación final para la imprenta (TUSTEP) o Internet (Collate). Sin embargo, diversos centros de investigación han ofrecido otros programas menos complicados, que pueden realizar la colación de manera sencilla y rápida, a la par que ofrecer otras informaciones y funcionalidades básicas.

En España, Marcos Marín desarrolló un programa llamado UNITE que era «capaz de comparar hasta 30 versiones de un mismo texto para obtener una versión única»¹⁷ (Admyte, 1991: 5). Este programa que se incluyó en el disco 0 de Admyte (1992-1998), tenía una fuerte limitación: estaba diseñado para trabajar con textos en verso. El único ejemplo de la aplicación de este programa es la edición unificada del *Libro de Alexandre* que él mismo publicó (1987)¹⁸.

Alonso Rioja (1996) desarrolló una aplicación para colación de textos que llamó AFTL, acrónimo de Análisis Filológico de Textos según el método de Lachmann. Lo diseñó para comparar «tres copias o traducciones de un mismo original» y «entresacar las diferencias existentes línea a línea y palabra a palabra entre los tres textos» (1996, 8). Con esta aplicación pretendía «automatizar el proceso de comparación siendo capaz de reconocer en un máximo de tres líneas consecutivas, todos los tipos de diferencias (omisión, interpolación, difracción o inversión)» (Alonso Rioja, 1996, 8). Aunque existen los discos, ha sido imposible instalar la aplicación de acuerdo con las instrucciones que se ofrecen en el manual de usuario de la aplicación AFTL (Alonso Rioja, 1996, 81-91). Sin embargo, se ofrece una serie de impresiones de las diversas ventanas que muestran qué era capaz de hacer la aplicación.

Un programa de colación básico, pero potente y flexible es Juxta. Se

¹⁷ La parte inglesa de este folleto es menos ambiciosa, pues indica que es «[a] program designed to collate and compare up to ten versions of a given poetic text in order to obtain their presumed archetype, serving thereby as an aid in the production of a critical edition of that text» (Admyte, s. d., 5).

¹⁸ Esta edición es hoy accesible a través de la Biblioteca Virtual Miguel de Cervantes <<http://www.cervantesvirtual.com/obra/libro-de-alexandre--0/>> (cons. 28/10/21) y la información técnica es, por tanto, accesible.

trata de una aplicación desarrollada en la University of Virginia¹⁹ escrita en Java que funciona en ordenadores Windows, Linux y Apple. Sin embargo, ha dejado de funcionar en las máquinas Apple cuyo sistema operativo sea 11.4 o superior, aunque sigue funcionando en el entorno Windows 10 y Linux. Es un software muy valioso dado que puede leer ficheros con etiquetado TEI.

El flujo de trabajo es que tras cargar los textos y seleccionar uno de ellos como término de comparación, tras unos segundos se oscurecen todas aquellas palabras en las que no hay igualdad. En una de las posibles visualizaciones, al posar el cursor sobre una palabra con variación se despliegan todas las variantes que presentan los diferentes testimonios (Fig. 2). Juxta tiene, incluso, la capacidad de generar el aparato crítico, que no es nada más que una página HTML que puede transformarse, con relativa sencillez, en un aparato crítico de acuerdo con el sistema TEI.

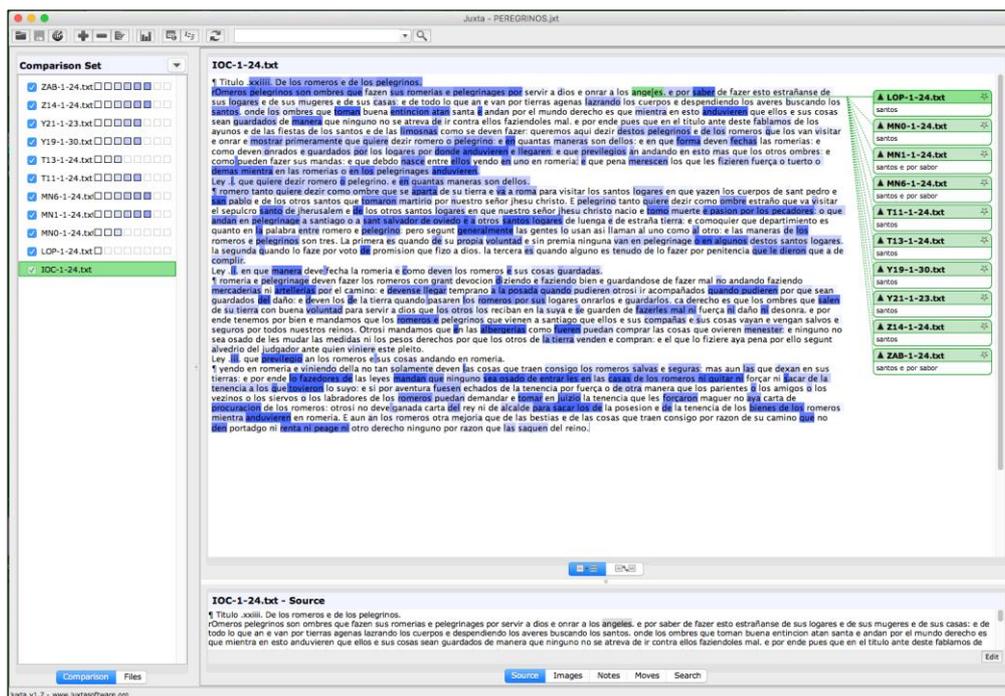


Fig. 2. Resultado en Juxta de la colación de todos los testimonios del título 1.24 de las *Siete Partidas*

¹⁹ <<https://www.juxtaoftware.org/>> (cons. 28/10/21).

Es un programa magnífico, pero tiene un problema: los ordenadores son obstinadamente literales y la variación ortográfica que presentan los manuscritos medievales provoca gran ruido. En una primera instancia se puede resolver el problema por medio de una regularización textual. En el caso de los textos castellanos medievales apliqué, aunque no comparto los criterios, los principios de presentación crítica desarrollados por Sánchez-Prieto Borja (1998; 2011). Para conseguirlo procesé el texto con la aplicación desarrollada para el etiquetado morfológico automático del Old Spanish Textual Archive (Gago Jover y Pueyo, 2018a; 2018b; 2020).

Puesto que Juxta no funciona en las máquinas Apple con sistemas operativos más recientes, se buscó otra solución para llevar a cabo la colación automática. La respuesta ha sido el programa CollateX²⁰. Aunque puede funcionar como una aplicación Java, es posible usarlo desde Python, lo cual me permite mecanizar ciertos aspectos del pre- y postprocesado de los textos.

```
from collatex import *
collation = Collation()
testimonio_IDI = open( "PEREGRINOS/IDI-1-24.txt", encoding='utf-8' ).read()
testimonio_IOC = open( "PEREGRINOS/IOC-1-24.txt", encoding='utf-8' ).read()
testimonio_LOP = open( "PEREGRINOS/LOP-1-24.txt", encoding='utf-8' ).read()
testimonio_MN0 = open( "PEREGRINOS/MN0-1-24.txt", encoding='utf-8' ).read()
testimonio_MN1 = open( "PEREGRINOS/MN1-1-24.txt", encoding='utf-8' ).read()
testimonio_MN6 = open( "PEREGRINOS/MN6-1-24.txt", encoding='utf-8' ).read()
testimonio_T11 = open( "PEREGRINOS/T11-1-24.txt", encoding='utf-8' ).read()
testimonio_T13 = open( "PEREGRINOS/T13-1-24.txt", encoding='utf-8' ).read()
testimonio_Y19 = open( "PEREGRINOS/Y19-1-30.txt", encoding='utf-8' ).read()
testimonio_Y21 = open( "PEREGRINOS/Y21-1-23.txt", encoding='utf-8' ).read()
testimonio_Z14 = open( "PEREGRINOS/Z14-1-24.txt", encoding='utf-8' ).read()
testimonio_ZAB = open( "PEREGRINOS/ZAB-1-24.txt", encoding='utf-8' ).read()
salida = open( 'pregrinos.txt', 'w', encoding='utf-8')
collation.add_plain_witness( "IOC", testimonio_IOC )
collation.add_plain_witness( "IDI", testimonio_IDI )
collation.add_plain_witness( "LOP", testimonio_LOP )
collation.add_plain_witness( "MN0", testimonio_MN0 )
collation.add_plain_witness( "MN1", testimonio_MN1 )
collation.add_plain_witness( "MN6", testimonio_MN6 )
collation.add_plain_witness( "T11", testimonio_T11 )
collation.add_plain_witness( "T13", testimonio_T13 )
collation.add_plain_witness( "Y19", testimonio_Y19 )
collation.add_plain_witness( "Y21", testimonio_Y21 )
collation.add_plain_witness( "Z14", testimonio_Z14 )
collation.add_plain_witness( "ZAB", testimonio_ZAB )
tabla = collate(collation, segmentation=False, near_match=True, layout='vertical')
print(tabla, file=salida)
```

Fig. 3. Script en Python para la colación de los testimonios

²⁰ <<https://collatex.net/>> (cons. 28/10/21).

Es muy sencillo de usar. La figura 3 muestra el *script* que he utilizado para colacionar el título 1.24, el último de la *Primera Partida*, en todos los testimonios a los que he tenido acceso. Las dos primeras líneas invocan la librería. Las líneas que comienzan con `testimonio_` leen el texto del título elegido en todos los testimonios que lo tienen. A continuación, se declara cómo se llamará el fichero de salida y qué codificación tendrá. En el siguiente paso, todas las líneas que comienzan con `collation_`, alimentan cada uno de los textos, pero identificados por la sigla, a la función `collation`, la cual tendrá en cuenta que no se quiere segmentar el texto, que se ha de aplicar el parámetro `near_match` y que el resultado ha de ofrecerse verticalmente, pues no sucede como en Juxta, en la que el resultado es visual.

CollateX puede ofrecer grafos (Figura 4) que permiten seguir la trayectoria y conexión entre las diferentes variantes y, aunque también puede ofrecer los resultados en TEI (según las normas del módulo *Critical Apparatus*), considero que la mejor opción es en forma de una tabla (Figura 5), pues en esta fase del proyecto porque permite detectar errores de transcripción que de otra manera pasarían desapercibidos. Reprocesar los ficheros tras las correcciones manuales no toma demasiado tiempo.

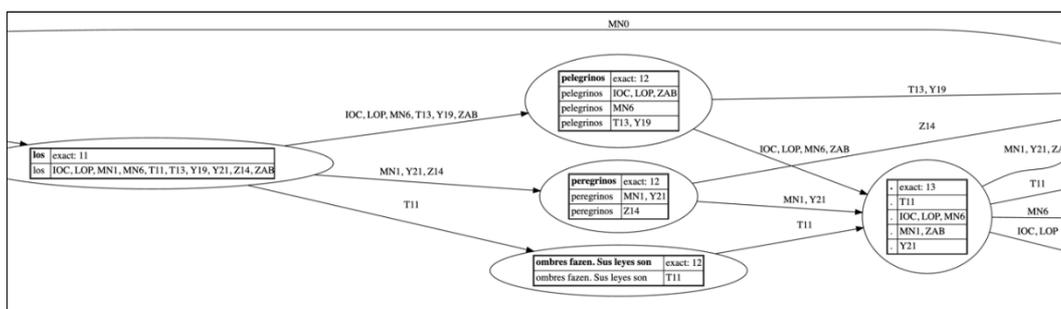


Fig. 4. Grafo (parcial) del resultado de la colación del título 1.24 de la *Primera Partida*

Antes he dicho que los ordenadores son recalcitrantemente literales y consideran como formas diferentes palabras como *rOmeros*, *Romeros*, *romeros*, lo que arroja un elevado número de variantes de nulo valor. Se puede reducir este ruido, como en el caso de Juxta, por medio de una transformación de los ficheros al modelo de presentación crítica

desarrollado por Sánchez-Prieto Borja (1998; 2011). Sin embargo, CollateX tiene integrado en su algoritmo la posibilidad de `near_match`, que, declarándolo como verdadero, obvia el problema y considera todas esas variantes como indiferentes, con lo que se reduce el ruido y las manipulaciones del texto.

| IOC | IDI | LOP | MN0 | MN1 | MN6 | T11 | T13 | Y19 | Y21 | Z14 | ZAB |
|------------|------------|------------|-----------------|--------------|------------|--------------|--------------|--------------|--------------|--------------|--------------|
| § | - | § | - | - | § | § | § | § | § | § | - |
| Título | Título | Título | Título | Título | Título | Título | Título | Título | Título | Título | Título |
| xxiiii | xxiiij | XXIIII | xxiiijº | xxiiii | xxiiij | xxiiijº | xxiiijº | xxxº | xxiiij | xxiiijº | xxvj |
| De | de | De | De | de | de | de | de | de | de | de | de |
| los | los | los | los | los | los | los | los | los | los | los | los |
| romeros | romeros | romeros | romeros | romeros | romeros | romerías | romeros | romeros | romeros | romeros | romeros |
| e | & | e | - | e | e | e | e | e | e | e | e |
| de | de | de | de | de | de | de | de | de | de | de | de |
| - | - | - | las | - | - | los | - | - | - | - | - |
| - | - | - | religiones | - | - | peregrinages | - | - | - | - | - |
| - | - | - | - | - | - | que | - | - | - | - | - |
| los | los | los | los | los | los | los | los | los | los | los | los |
| pelegrinos | pelegrinos | pelegrinos | - | peregrinos | pelegrinos | ombres | pelegrinos | pelegrinos | peregrinos | peregrinos | pelegrinos |
| rOmeros | RÓmeros | rOmeros | romerías | romerías | romeros | fazen | romerías | romerías | romerías | romerías | romerías |
| - | - | e | e | e | e | Sus | - | - | - | - | - |
| pelegrinos | pelegrinos | pelegrinos | peregrinaciones | - | pelegrinos | leyes | - | - | - | - | - |
| son | son | son | - | - | - | son | - | - | - | - | - |
| ombres | ombres | ombres | - | - | - | iiijº | - | - | - | - | - |
| - | - | - | - | - | - | Romerías | - | - | - | - | - |
| que | que | que | - | - | se | e | e | e | e | e | e |
| - | - | - | - | peregrinages | - | peregrinages | pelegrinages | pelegrinages | peregrinages | peregrinages | pelegrinages |
| fazen | fazen | fazen | fazen | fazen | fazen | fazen | fazen | fazen | fazen | fazen | fazen |
| sus | sus | sus | los | los | los | los | los | los | los | los | los |
| romerías | romerías | romerías | ombres | ombres | ombres | ombres | ombres | ombres | ombres | ombres | ombres |
| e | & | e | - | - | - | - | - | - | - | - | - |

Fig. 5. Tabla que ofrece como resultado el *script* de la figura 3

La tabla (Fig. 5) que ofrece el *script* es un tanto muda. Tras editarla para convertirla en una hoja de Excel (Fig. 6) es mucho más elocuente, dado que en Excel tiene la posibilidad de marcar las filas (casillas) en las que hay una diferencia. Aquí volvemos a la literalidad de los ordenadores, aunque no es tan desesperante como en el caso de Juxta porque Excel no considera como diferentes *rOmeros*, *Rómeros*, *romeros* (línea 15) pero sí la ausencia de un término (línea 22), o el cambio de un término o forma por otro: *las* por *los* (línea 6), *romeros* frente a *romerías* (línea 7), o *romerías* frente a *ombres* (línea 25).

| 1 | IOC | IDI | LOP | MNO | MN1 | MN6 | T11 | T13 | Y19 | Y21 | Z14 | ZAB |
|----|------------|------------|------------|---------------------|------------|--------------|---------------------|---------------------|------------------|--------------|---------------------|--------------|
| 2 | ¶ | ¶ | ¶ | ¶ | ¶ | ¶ | ¶ | ¶ | ¶ | ¶ | ¶ | ¶ |
| 3 | Titulo | Titulo | Titulo | Titulo | Titulo | Titulo | Titulo | Titulo | Titulo | Titulo | Titulo | Titulo |
| 4 | xxiii | xxiiij | XXIII | xxiiij ^a | xxiii | xxiiij | xxiiij ^a | xxiiij ^a | xxx ^a | xxiiij | xxiiij ^a | xxvj |
| 5 | De | De | De | De | de | de | de | de | de | de | de | de |
| 6 | los | los | los | los | los | los | las | los | los | los | los | los |
| 7 | romeros | romeros | romeros | romeros | romeros | romeros | romerias | romeros | romeros | romeros | romeros | romeros |
| 8 | e | & | e | - | e | e | e | e | e | e | e | e |
| 9 | de | de | de | de | de | de | de | de | de | de | de | de |
| 10 | - | - | - | las | - | - | los | - | - | - | - | - |
| 11 | - | - | - | religiones | - | - | peregrinages | - | - | - | - | - |
| 12 | - | - | - | - | - | - | que | - | - | - | - | - |
| 13 | los | los | los | - | los | los | los | los | los | los | los | los |
| 14 | peregrinos | peregrinos | peregrinos | - | peregrinos | peregrinos | ombres | peregrinos | peregrinos | peregrinos | peregrinos | peregrinos |
| 15 | romeros | romeros | romeros | romerias | romerias | romeros | fazen | romerias | romerias | romerias | romerias | romerias |
| 16 | - | - | e | e | e | e | Sus | - | - | - | - | - |
| 17 | peregrinos | peregrinos | peregrinos | peregrinaciones | - | peregrinos | leyes | - | - | - | - | - |
| 18 | son | son | son | - | - | - | son | - | - | - | - | - |
| 19 | ombres | ombres | ombres | - | - | - | lijij ^a | - | - | - | - | - |
| 20 | - | - | - | - | - | - | Romerias | - | - | - | - | - |
| 21 | que | que | que | - | - | se | e | e | e | e | e | e |
| 22 | - | - | - | peregrinages | - | peregrinages | peregrinages | peregrinages | peregrinages | peregrinages | peregrinages | peregrinages |
| 23 | fazen | fazen | fazen | fazen | fazen | fazen | fazen | fazen | fazen | fazen | fazen | fazen |
| 24 | sus | sus | sus | los | los | los | los | los | los | los | los | los |
| 25 | romerias | romerias | romerias | ombres | ombres | ombres | ombres | ombres | ombres | ombres | ombres | ombres |
| 26 | e | & | e | - | - | - | - | - | - | - | - | - |

Fig. 6. La misma tabla de la figura 5 pero editada para manejarla en Excel

El uso de estos métodos informáticos me llevó a detectar, por ejemplo, problemas lingüísticos que de otra manera habrían pasado desapercibidos. En un recuento léxico por lemas de las ediciones *princeps* y de 1555 de las *Siete Partidas* constaté una interesante diferencia: la edición del quinientos tiene, en la *Primera Partida*, un tinte lingüístico medieval que la de 1491 no presenta. Simplificando la cuestión, en el texto de López hay preferencia por *maguer*, *guisar*, *toller*, *vegadas* y *ca* con una diferencia de uso que oscila entre el 100 % (caso de *toller*) y el 50 % (*guisa*), frente a las formas del texto de 1491. Esto llevó a preguntarnos cómo un texto que, sabemos, es copia de otro anterior (Fradejas Rueda, 2021c), pudo dar un paso atrás lingüísticamente.

He llegado a la respuesta por medio de la bibliografía material (Fradejas Rueda, 2021a), pero la respuesta incontestable la ofrecerán los métodos digitales. Ahí se ha incorporado una nueva herramienta al flujo de trabajo de 7PartidasDigital: la transcripción automática por medio de Transkribus²¹. Ya la he probado con la edición de diciembre de 1491 de las *Siete Partidas*, gracias al modelo desarrollado para las góticas del siglo XV y XVI castellanas (Bazzaco, 2020). Con este modelo he obtenido un porcentaje de acierto del 99.3 %, o lo que es lo mismo, solo el 0.7 % de las formas son erróneas, por lo que se han transcrito las casi mil páginas que constituyen esta edición en poco menos de una semana²².

²¹ <<https://readcoop.eu/transkribus/>> (cons. 28.10.21).

²² Eduardo Camero Santos, estudiante de doctorado en la Universidad de Valladolid, va a poner el sistema a prueba con la transcripción y codificación de la edición de 1528 de las *Siete Partidas*, pues,

Transkribus puede entregar un fichero docx, un fichero pdf, un fichero en texto plano, pero también un fichero etiquetado en TEI; sin embargo, es un etiquetado que solo se preocupa de los aspectos estructurales básicos: encabezados, columnas, líneas, además de los referentes a las imágenes que en 7PartidasDigital no son de interés. Por eso se ha desarrollado un *script* en R²³ para transformar el fichero TEI de Transkribus y adecuarlo al modelo editorial establecido en el proyecto.

Lo importante, y con esto finalizo, es que los viejos y probadísimos métodos de la filología clásica se pueden y tienen que complementar con los métodos y herramientas digitales que tenemos al alcance de la mano. Además, no podemos contentarnos con las herramientas cerradas que nos ofrecen la mayoría de los programas, tenemos que complementarlos con el manejo de un lenguaje de programación que ayude a resolver, en segundos, problemas mecánicos y repetitivos que manualmente llevarían muchísimas horas de trabajo y con el riesgo de errores. También hay que recordar, y tener en cuenta, que debemos mantener los ficheros de las transcripciones lo más puros y limpios posible: GitHub²⁴ y su control de versiones es la mejor opción en la actualidad.

§

como se ha demostrado (Fradejas Rueda, en prensa), en esta edición cuidada por Francisco de Velasco, se introduce la medievalización del texto que muestra Gregorio López. Su objetivo final es determinar qué corrigió Gregorio López, qué corrigió Francisco de Velasco y hasta qué punto cambia la lengua de una edición a otra.

²³ Véase: <<https://github.com/7PartidasDigital/XML-TEI/blob/master/scripts/Transkribus-7PD.R>> (cons. 27/10/2021).

²⁴ 7PartidasDigital guarda sus ficheros TEI en <<https://github.com/7PartidasDigital/XML-TEI>> (Fradejas Rueda, 2018).

Bibliografía citada

- Admyte, *Admyte*. Madrid: Micronet, Quinto Centenario, Biblioteca Nacional, 1991.
- Admyte, *Archivo digital de manuscritos y textos españoles [ADMYTE]* (1992-1998), eds. Francisco Marcos Marín, Gerardo Meiro, Charles B. Faulhaber, Ángel Gómez Moreno, Aurora Martín de Santa Olalla, Julián Martín Abad y John Nitti, Madrid, Micronet, 1992-1998, vols. 0, 1 y 2.
- Alonso Rioja, Valvanera, «Herramienta para análisis filológico según el método de Lachman (AFTL)», Valladolid, Universidad, Proyecto Fin de Carrera, 1996.
- Bazzaco, Stefano, «El reconocimiento automático de textos en letra gótica del Siglo de Oro: creación de un modelo HTR basado en libros de caballerías del siglo XVI en la plataforma Transkribus», *Janus*, 9 (2020), 534-561 <<https://www.janusdigital.es/articulo.htm?id=160>> (cons. 29/10/2021).
- Bleuca, Alberto, *Manual de crítica textual*, Madrid, Castalia, 1983.
- Buelow, Kenneth y David Mackenzie, *A Manual of Manuscript Transcription for the Dictionary of the Old Spanish Language*, Madison, HSMS, 1977.
- Cárdenas, Anthony J., John J. Nitti y Jean Gilkison Mackenzie, *Bibliography of Old Spanish Texts (literary texts)*, Madison, Hispanic Seminary of Medieval Studies, 1975.
- Castrillo Benito, Nicolás, *Programas Tustep (TUEbinger System von TExtverarbeitungProgrammen): aplicación del tratamiento de textos a la investigación*, Valladolid, Universidad de Valladolid, 1992.
- Castro, Américo, «La crítica filológica de los textos», *Boletín de la Institución Libre de Enseñanza*, 41, n. 682 (1917), 26-31.
- , «La crítica filológica de los textos», *Lengua, enseñanza y literatura (esbozos)*, Madrid, Victoriano Suárez, 1924, 171-197.
- Craddock, Jerry R., «La nota cronológica inserta en el prólogo de las *Siete Partidas*», *Al-Andalus*, 39 (1924), 363-390.

- Dearing, Vinton A., *Methods of Textual Editing. A Paper Delivered at a Seminar on Bibliography Held at the Clark Library, 12 May, 1962*, Los Angeles, William Andrews Clark Memorial Library-University of California, 1962.
- Faulhaber, Charles B., «La *Text Encoding Initiative* y su aplicación a la codificación textual y explotación», en *Actas del Congreso de la Lengua Española: Sevilla, 7 al 10 octubre, 1992*, Madrid, Instituto Cervantes, 1994, 331-340.
- , *PhiloBiblon*, Bancroft Library, University of California, Berkeley, 1997, <<http://vm136.lib.berkeley.edu/BANC/philobiblon/index.html>> (cons. 28.10.2021).
- Faulhaber, Charles B. y Ángel Gómez Moreno, *Normas para BOOST4 (Bibliography of Old Spanish Text 4th Edition)*, Madison, HSMS, 1986.
- Faulhaber, Charles, Ángel Gómez Moreno, Anthony J. Cárdenas, John J. Nitti, y Jean Gilkison Mackenzie, *Bibliography of Old Spanish Texts* (3.^a edición), Madison, Hispanic Seminary of Medieval Studies, 1984.
- Fernández-Ordóñez, Inés, «Reseña a “Alberto Blecua: Manual de crítica textual, Madrid: Castalia, 1983”», *Edad de Oro*, 7 (1988), 231-240.
- Foulet, Alfred y Mary Blakely Speer, *On Editing Old French Texts*, Lawrence, The Regents Press of Kansas, 1979.
- Fradejas Lebrero, José y José Manuel Fradejas Rueda, *Pero López de Ayala, Libro de la caza de las aves*, Barcelona, Castalia, 2016.
- Fradejas Rueda, José Manuel, «*Tratado de cetrería*. Texto, gramática y vocabulario (según el Ms. 9 de la R.A.E.)», Madrid, Universidad Complutense, 2 vols., 1983.
- , *Texto y concordancias de los textos menores del MS. V.II.19 de El Escorial: «Gerardus falconarius», «Dancus Rex», «Guillelmus falconarius», «Libro de los açores»*, Madison, HSMS, 1992a.
- , *Texto y concordancias del MS Additional 16392 de la British Library: «Libro de la caza de las aves» de Pero López de Ayala*, Madison, HSMS, 1992b.
- , *Historia de Enrique, fi de Oliva*, Madrid, Centro Virtual Cervantes, 1997 <<https://cvc.cervantes.es/literatura/clasicos/fi/default.htm>> (cons. 28/10/21).
- , *Textos clásicos de cetrería, montería y caza*, Madrid, Fundación Histórica Tavera, Digibis, 1999.

- , «La Codificación XML/TEI de Textos Medievales», *Memorabilia*, 12 (2010), 219-247 <<http://parnaseo.uv.es/Memorabilia/Memorabilia12/PDFs/Codificacion.pdf>> (cons. 28/10/21).
 - , *La versión castellana medieval de la «Epitome rei militaris»*, San Millán de la Cogolla, Cilengua, 2011.
 - , «Cuatro nuevos testimonios manuscritos de las *Siete Partidas*», *Revista de literatura medieval*, 27 (2015), 13-52.
 - , 7PartidasDigital/XML-TEI: Primera versión (Version v.0.1), Zenodo, 2018 <<http://doi.org/10.5281/zenodo.1195642>> (cons. 28/10/21).
 - , «Incunables de las *Siete Partidas* en Hispanoamérica» , en *Las “Siete Partidas” del Rey Sabio una aproximación desde la filología digital y material*, eds. J. M. Fradejas, E. Jerez y R. Pichel Madrid, Iberoamericana, 2021a, 175-189.
 - , «La codificación TEI de las ediciones de 1491 y 1555 de las *Siete Partidas*», en *Las “Siete Partidas” del Rey Sabio: Una aproximación desde la filología digital y material*, eds. J. M. Fradejas Rueda, E. Jerez Cabrero y R. Pichel, Madrid, Iberoamericana, 2021b, 253-265.
 - , «Las *Siete Partidas*: del pergamino a la red», en *Alfonso el Sabio y la conceptualización jurídica de la monarquía en las “Siete Partidas”*, eds. Mechthild Albert, Ulrike Becker y Elmat Schmidt, Bonn, University Press, 2021c, 223-264.
 - , «Los testimonios castellanos de las *Siete Partidas*», en *Las “Siete Partidas” del Rey Sabio: Una aproximación desde la filología digital y material*, eds. J. M. Fradejas Rueda, E. Jerez Cabrero y R. Pichel, Madrid, Iberoamericana, 2021d, 21-35.
 - , «Las *Siete Partidas*: el texto base de la edición de Gregorio López (1555)», *Actas del Congreso Historiador y Poder, el Historiador en el Poder. VIII Centenario del nacimiento de Alfonso X el Sabio*, Moscú, en prensa.
- Gago Jover, Francisco y F. Javier Pueyo Mena, «El *Old Spanish Textual Archive*, diseño y desarrollo de un corpus de textos medievales: lematización y etiquetado gramatical», *Scriptum Digital*, 7 (2018a), 25-35.

- , «El *Old Spanish Textual Archive*, diseño y desarrollo de un corpus de textos medievales: el corpus textual», *Cuadernos del Instituto Historia de la Lengua*, 11 (2018b), 165-209.
- , *Old Spanish Textual Archive*. Hispanic Seminary of Medieval Studies 2020 <<http://osta.oldspanishtextualarchive.org>> (cons. 29/10/21).
- García Solalinde, Alfonso, *Alfonso X el Sabio, General Estoria. Primera Parte*. Madrid, Centro de Estudios Históricos, 1930.
- Herrero de la Fuente, Marta, «*Alma littera*»: *Estudios dedicados al profesor José Manuel Ruiz Asencio*, Valladolid, Universidad, 2014.
- Hockey, Susan, *A Guide to Computer Applications in the Humanities*, London, Duckworth, 1980.
- , *Electronic Texts in the Humanities: Principles and Practice*, Oxford, OUP, 2000.
- Jauralde Pou, Tablo, *Manual de investigación literaria*, Madrid, Gredos, 1981.
- Kasten, Lloyd A., John J. Nitti y Wilhelmina Jonxis-Henkemans, *The Electronic Texts and Concordances of the Prose Works of Alfonso X, El Sabio*, Madison, HSMS, 1997.
- Lapesa, Rafael, *Canciller Pero López de Ayala. Rimado de Palacio: esbozo de una edición crítica*, Valencia, Biblioteca Valenciana, 2010.
- Lázaro Carreter, Fernando, *Diccionario de términos filológicos*, Madrid, Gredos, 1952.
- López Estada, Francisco, *Introducción a la literatura medieval española*, Madrid, Gredos, 1979.
- Manley, John M. y Edith Rickert, *The Text of the Canterbury Tales*, Chicago, U. Chicago Press, 1940.
- Manuscrito = *Manuscrito de Per Abbat. Cantar de Mio Cid*, Madrid, Biblioteca Nacional de España, 1998.
- Marcos Marín, Francisco, *Libro de Alexandre*, Madrid, Alianza Editorial, 1987.
- Marín Ocete, A., «El estado actual de la crítica de textos», *Boletín de la Universidad de Granada*, 4 (1932), 349-361.
- Moorman, Charles, *Editing the Middle English Manuscript*, Jackson, U.P. of Mississippi, 1975.
- Nitti, John, «Computers and the Old Spanish Dictionary», *Computers and the Humanities*, 12 (1978), 43-52.

- O'Neill, John, *Electronic Text and Concordances of the Madison Corpus of Early Spanish Manuscripts and Printings*, Madison-New York, HSMS, 1999.
- Onís, Federico, *Torres Villarroel. Vida*, Madrid, La Lectura, 1912.
- Robinson, Peter M. W., «The Collation and Textual Criticism of Icelandic Manuscripts (1): Collation», *Literary and Linguistic Computing*, 4/2 (1989a), 99–105.
- , «The Collation and Textual Criticism of Icelandic Manuscripts (2): Textual Criticism», *Literary and Linguistic Computing*, 4/3 (1989b) 174–181.
- Ruiz Albi, Irene, «Un fragmento de los *Bocados de oro* en el archivo de la Real Chancillería de Valladolid», en «*Alma littera*»: *Estudios dedicados al profesor José Manuel Ruiz Asencio*, ed. Herrero de la Fuente, Valladolid, Universidad, 2014, 579-593.
- Ruiz, Elisa (1985), «Crítica textual, edición de textos», en *Métodos de estudio de la obra literaria*, ed. J. M.^a Díez Borque, Madrid, Taurus, 67-120.
- , *Manual de codicología*, Salamanca, Fundación Germán Sánchez Ruipérez, 1988.
- Sánchez-Prieto Borja, Pedro, *Cómo editar los textos medievales. Criterios para su presentación gráfica*. Madrid, Arco/Libros, 1998.
- , *La edición de textos españoles medievales y clásicos. Criterios de presentación gráfica*, San Millán de la Cogolla, Cilengua, 2011.
- Santiago, Ramón, «Crítica textual y edición de textos en el *Diccionario de términos filológicos*: la primera descripción del método lachmanniano en España», en *Palabras, norma, discurso. En memoria de Fernando Lázaro Carreter*, Salamanca, Universidad, 2005, 1105-1119.
- , «Acerca de los primeros pasos de la crítica textual en la filología española: Menéndez Pidal y el Centro de Estudios Históricos», en *El legado de Ramón Menéndez Pidal (1869-1968) a principios del siglo XXI*, ed. Inés Fernández-Ordóñez, Madrid, CSIC, I, 2020, 221-256.
- Simón Díaz, José, *Bibliografía de la literatura hispánica*, Madrid, CSIC, 1950-1993, 16 vols.
- Wehrli, Max, *Introducción a la ciencia literaria*, Buenos Aires, Nova, 1951.
- West, Martin L., *Textual Criticism and Editorial Technique Applicable to Greek and Latin Texts*, Stuttgart, B.G. Teubner, 1973.