

ÍNDICE

<i>Nerea Parro Gómez</i> SOBRE LAS HUMANIDADES DIGITALES EN LA FILOLOGÍA HISPÁNICA	9
<i>Elena Cano Turrión</i> DESAFÍOS ACTUALES DE LAS HUMANIDADES DIGITALES Y EL ETIQUETADO TEI EN LOS CORPUS LITERARIOS ÁUREOS. LOS PROYECTOS SILEM, PRESOLO Y PARANOBA	27
<i>Soledad Castaño Santos</i> DIFICULTADES Y RETOS EN LA ELABORACIÓN DE PROYECTOS DE HUMANIDADES DIGITALES: EL PROYECTO “MAPPING THE SCRIPTURES IN WESTERN SEPHARDI LITERATURE”	53
<i>Juan Esteban Hincapié Atehortúa</i> DE LA FILOLOGÍA DEL TEXTO IMPRESO A LAS HUMANIDADES DIGITALES: FRUTOS- DIGITAL, EDICIÓN CRÍTICO-GENÉTICA DIGITAL DE <i>FRUTOS DE MI TIERRA</i> DE TOMÁS CARRASQUILLA	83
<i>Miguel Calderón Campos</i> LA HISTORIA DEL DIMINUTIVO -ICO EN EL ESPAÑOL PENINSULAR (1500-1799)	121
<i>Inmaculada González Sopena y Gael Vaamonde</i> ANOTACIÓN Y ANÁLISIS PARA EL ESTUDIO DIACRÓNICO DE ARABISMOS EN ORALIA DIACRÓNICA DEL ESPAÑOL (ODE).....	153
<i>Andrés Echavarría Peláez y Bárbara Santiago Medina</i> LAS NORMAS DEL SECRETO: EL <i>ORDEN DE PROCESSAR</i> DE PABLO GARCÍA Y LA TEI COMO MODELOS DE ESTANDARIZACIÓN DOCUMENTAL.....	193
SOBRE LOS AUTORES	241

SOBRE LAS HUMANIDADES DIGITALES EN LA FILOLOGÍA HISPÁNICA

Nerea Parro Gómez
Université de Neuchâtel

Que las humanidades digitales irrumpieron con fuerza hace años en el ámbito de la filología es un hecho indiscutible. Es numerosa la bibliografía que ahonda en la historia del uso de lo digital en el campo de las humanidades desde sus orígenes —que la mayoría de los autores establece en el proyecto¹ del jesuita Roberto Busa para la indexación de la obra completa de Tomás de Aquino con el apoyo económico y la tecnología de IBM, que dio lugar a la publicación del *Index Thomisticus* (Busa 1980; 2022 [1962])²—. Sin embargo, no es el propósito de estas páginas profundizar en la historia y evolución de las humanidades digitales ni detallar sus hitos (y mitos), sino ofrecer un panorama amplio de su campo de acción en lo que respecta a la labor del filólogo. Ahora bien, los expertos tardaron un tiempo en decidir el término que daría nombre a esta disciplina. Así se explica en la introducción al volumen titulado *Defining Digital Humanities*:

Over the past years, the field that we now refer to as digital humanities has been known by many terms: humanities computing, humanist informatics, literary and linguistic computing and digital resources in the humanities, to name but a few. Most recently it has predominantly been known as digital humanities, though other variations such as eHumanities are occasionally to be found in literature emanating from continental Europe (Terras, Nyhan y Vanhoutte 2013: 2).

¹ Aunque la mayoría de los autores afirman que este es el primer proyecto en humanidades digitales, muchos apuntan como precursora a Ada Lovelace en el siglo XIX (Hernández-Lorenzo 2022: 153).

² Un proyecto llevado a cabo por Alarcón (2002) y Alarcón y Bernot (2005) ha sido el convertir este trabajo del padre Busa en un corpus al uso en la actualidad. Puede consultarse en: <<https://www.corpusthomicum.org/it/index.age>>.

Si bien es cierto que en determinadas ocasiones los términos para el mundo anglófono no tienen necesariamente el mismo significado cuando se usan para el ámbito hispánico —Del Rio Riande (2014: 33; 2015: 15-17) insiste en la dificultad que supone transponer esta disciplina heredada de los *English Departments* (cf. Kirschenbaum 2010) a las humanidades digitales hispánicas y Priani Saisó *et al.* (2014: 11) observan una diferencia de uso entre el término en español y en inglés—, en este caso, también se ha reflexionado sobre cómo denominar esta disciplina para el español. En esta misma línea, Galina Russell (2011: 3) ha observado los siguientes términos “Recursos Digitales para las Humanidades, Cómputo para las Humanidades, Cómputo en las Humanidades, Informática Digital y Cultural e Informática para las Humanidades”, aunque a lo largo de la bibliografía publicada puede advertirse que se utilizan también otros para referirse a este tema como ‘recursos digitales’, ‘nuevas tecnologías’, ‘filología digital’, ‘informática para las humanidades’, ‘informática humanística’, ‘metodología informática’, ‘*humanities computing*’ y demás términos que algunos autores prefieren mantener en inglés, hasta decidirse por el sintagma ‘humanidades digitales’ o ‘*Digital Humanities*’³.

Ahora bien, lo complicado en esta disciplina —o línea de trabajo, etiqueta, metodología o campo (Rockwell 2013 [1999]: 13-33; Rodríguez-Yunta 2013; Del Rio Riande 2015: 17), porque parece que también hay dificultad para determinar su naturaleza ontológica— no fue cómo denominarla, ya que observábamos más arriba cómo la multiplicidad de términos desembocó en ‘humanidades digitales’ o ‘*Digital Humanities*’. Lo realmente complicado para los expertos parece ser el hecho de proponer una definición que concrete qué son. En general, las propuestas de los autores se caracterizan por la cautela a la hora de dar una definición de lo que son las humanidades digitales —e incluso de lo que no son—, aunque, a veces, siguen siendo poco concretas y demasiado vagas, quizá debido a que se sitúan en medio de muchas disciplinas.

Un ejemplo de la disparidad de definiciones podría ilustrarse mediante la existencia de una web (imagen 1) que recoge algo más de ochocientas definiciones de lo que diferentes expertos entienden por humanidades digitales o *Digital Humanities*. Estas citas, unas más extensas y elaboradas y otras más

³ Kirschenbaum (2010: 56-57) propone el origen de esta denominación en una discusión entre los editores sobre el título que le darían al volumen que acabó llamándose *A Companion to Digital Humanities*, publicado en 2004.

breves y concretas, fueron recogidas entre 2009 y 2014 con motivo del Día de las Humanidades Digitales.

What Is Digital Humanities?

New quote

DH is the perfect cocktail of computer science and rationalization of culture and humanities. Elena González-Blanco

NB: Click the [New quote](#) button to display a new random quote. Quotes were pulled from participants from the [Day of DH](#) between 2009-2014. As of January 2015, the database contains 817 rows and randomly selects a quote each time the page is loaded. If you want to do something cool with the data, I am providing the code and data [here](#).

Made by [Jason Heppler](#). Problems? Questions? I'm [@jaheppler](#) on Twitter.

Imagen 1⁴

Entrada de la página web “What is Digital Humanities?” (Heppler 2015)

Centrado en el ámbito hispánico, Del Rio Riande proporciona toda una serie de definiciones aportadas por diferentes hispanistas de ambos lados del Atlántico (2014: 21-25), todas en la línea de la enunciada por Rojas Castro (2013a: 79), que fue una de las primeras propuestas acerca de lo que se entiende por humanidades digitales para este ámbito, quien las describe como un “conjunto de principios, valores y prácticas en donde convergen múltiples objetos de estudio y saberes cuyas fronteras se encuentran en continua negociación”. Con todo, tal vez lo relevante no sea insistir en establecer una definición exacta, sino enfocarse en identificar cuáles son las características comunes que se repiten a lo largo de las diferentes descripciones. Una propuesta parecida a esta idea es la de Hernández-Lorenzo (2022: 142-152), dado que destaca algunas de las características que también están presentes en la mayoría de definiciones, como: la interdisciplinaridad; el trabajo colaborativo entre especialistas de distintos ámbitos de estudio; la internacionalización; la difusión en acceso abierto de códigos y métodos; la institucionalización gracias a diferentes organismos, centros académicos o laboratorios; o el trabajo

⁴ Se ha escogido una definición al azar de las ochocientas diecisiete disponibles en mayo de 2025, cuya autora es Elena González-Blanco.

con grandes cantidades de datos también llamados *big data*. De lo anterior se desprende, sin querer aportar una nueva definición, sino simplemente recoger los puntos en común orientados al campo de la filología —ya que ese es el propósito de este volumen—, que se trata del punto de encuentro del saber filológico con la metodología y tecnología digitales que daría como resultado nuevas herramientas para la edición y estudio de fuentes para la investigación en lengua y literatura y otros recursos útiles para el análisis y presentación de datos que, además, pueden compartirse con toda la comunidad investigadora y pueden difundirse en acceso abierto.

Aunque cada vez menos, esta nueva proyección de los estudios filológicos puede suponer un cambio de mentalidad que parecería chocar con el concepto tradicional del filólogo que se encierra solitario en una biblioteca, rodeado de libros y fichas e investigando de manera aislada en sus manuscritos (González-Blanco 2013). Ese filólogo mostraría ahora otra faceta como filólogo digital, que ha adquirido formación tecnológica e informática, cuyo trabajo es individual y a la vez colaborativo, ya que tras la edición de textos o creación de corpus textuales —partiendo de la base de que el texto es una de las fuentes principales de estudio del filólogo— suele haber todo un equipo trabajando en conjunto, y que, además, está conectado con otros investigadores para poder compartir los métodos de trabajo y resultados. Eso sí, la única manera de poder llevar a cabo este tipo de proyectos sería contando con una sólida formación filológica: en palabras de Fradejas (2022: 40) “*primum philologia, deinde computatrum scientia*, no se puede ser editor digital sin una sólida base de filología y crítica textual”. En efecto, la figura del filólogo tradicional se revela como un componente esencial en la conformación del filólogo digital. En este sentido, solo aquel que es filólogo de manera “analógica”, también puede serlo en el mundo digital, toda vez que las decisiones sobre el objeto de estudio deben ser tomadas desde una formación consolidada que garantice el rigor científico en el paso desde el papel hasta el formato digital.

Si bien este escenario representa el ideal del quehacer filológico unido a las posibilidades que ofrece el mundo digital, uno de los desafíos actuales es precisamente la formación tecnológica básica de los investigadores. En la mayoría de los casos, el autoaprendizaje ha supuesto la única manera para la adquisición de competencias digitales tales como conocimientos en programación, herramientas de análisis, procesamiento y presentación o visualización de datos, etiquetado de textos, servidores web, creación y gestión de bases de datos, etc. Precisamente porque esta formación tecnológica se ve algo

limitada, es habitual formar equipo junto a técnicos o ingenieros informáticos. Esta simbiosis de ciencias y letras es, en muchas ocasiones, la única combinación que hace posible la realización del proyecto pensado primeramente por el filólogo y que se plasma en el mundo digital gracias al ingeniero. A pesar de esto, la formación tecnológica de los filólogos es cada vez más completa y especializada y la formación en humanidades digitales se va implementando paulatinamente en los planes de estudios y cursos destinados a investigadores.

Retomando una de las ideas anteriores, el texto es una de las fuentes principales del filólogo como objeto de estudio, y por esta razón, no es de extrañar que también fuera el objeto de los primeros proyectos que se desarrollaron desde la implementación de lo digital en el campo de la filología hispánica, que comenzó su andadura a finales de los años 80 (Rojas Castro 2013b; Spence y González-Blanco 2014). Así, a partir de esta época comenzó a publicarse bibliografía sobre este tema⁵ y, unos años más tarde, empezaron a digitalizarse archivos y bibliotecas al mismo tiempo que se creaban repositorios digitales de textos que ya nacían digitales⁶. Estos tipos de edición digital podrían corresponderse con las dos primeras tipologías de edición digital descritas por Pierazzo (2019a: 179) —también por Torruella y Llisterri (1999: 51), aunque enfocadas únicamente hacia la lingüística de corpus—: el *archivio editorializzato* o “archivo/colección informatizado” (podría ser, según lo descrito por los autores, una colección de imágenes digitales, provista de metadatos y que puede contar con una catalogación del contenido) y la *edizione di lettura* o “Biblioteca de Textos Electrónicos” para Pierazzo (2019a: 179) tendría el énfasis en la simplicidad de acceso al texto y la fiabilidad científica ya que, para este tipo de ediciones normalmente se dispone de las imágenes facsimilares o de una transcripción con criterios explícitos y notas al texto⁷, mientras que para Torruella y Llisterri (1999: 51) esta colección de textos no seguiría un riguroso criterio de selección. Sin embargo, el objeto de estudio iría hoy más allá de la digitalización para la difusión y acceso abierto de ciertas obras, lo que nos

⁵ Por ejemplo, uno de los primeros libros publicados fue el de Marcos Marín en 1994 y un artículo de Lucía Megías en 1998.

⁶ Podemos destacar la digitalización de la Biblioteca Nacional de España o de archivos históricos nacionales y provinciales y otros proyectos nacidos digitales como el Archivo Digital de Manuscritos y Textos Españoles (ADMYTE), la Biblioteca Virtual Miguel de Cervantes, el Corpus de Referencia del Español Actual (CREA) y el Corpus Diacrónico del Español (CORDE), entre muchos otros.

⁷ Adaptación del original en italiano (Pierazzo 2019a: 179-180).

llevaría al tercer tipo que la autora distingue: la *edizione arricchita*. Esta “edición enriquecida” sería la manera en la que se hace hincapié en la calidad del paratexto, la posibilidad de realizar diferentes presentaciones del mismo texto (paleográfica, crítica, diplomática, facsimilar, etc.) y puede ir acompañado de anotaciones filológicas, históricas, de un aparato crítico, etc. (Pierazzo 2019a: 180). Resulta interesante cómo la propuesta de esta autora intenta establecer tres tipologías generales que serían clasificables sea cual sea el acercamiento metodológico o teórico del investigador al preparar la edición⁸. En este sentido, parece esencial establecer unos criterios claros que permitan separar de manera explícita la edición del propio texto de los añadidos del editor. Es precisamente la edición enriquecida —o corpus enriquecido—, que es la que se aplica generalmente a los textos literarios y documentos históricos, lo que le interesa generalmente al filólogo, dado que esta implica la adición de metadatos, marcado estructural y anotaciones lingüísticas que incrementan su valor para la investigación (búsquedas, análisis, presentación de datos, etc.) y la conservación digital.

Si se tiene en consideración el marco anterior, un corpus lingüístico o literario podría explicarse —de manera sumamente simplificada ya que para cada una de las ramas de esta disciplina existirían diferencias ligadas a los intereses del estudio— como una colección de textos debidamente recopilados en función de unos criterios marcados por el objeto de estudio. Sería pertinente realizar una aclaración, y es que un corpus no necesariamente debe cumplir el requisito de estar alojado en un soporte digital, pues siguen existiendo y teniendo gran utilidad los corpus realizados de manera tradicional:

[...] el término corpus se relaciona casi indefectiblemente con lo digital [...]. Elaboración de corpus e informática han confluído, pero no siempre se han identificado, pues hay corpus elaborados para edición en papel. Es evidente, sin embargo, que ahora no pueden concebirse por separado (Sánchez-Prieto Borja 2012: 29).

Sin embargo, sigue existiendo un problema y es que, aunque se marquen los criterios que ha seguido el proceso de edición de un texto, ¿cómo puede saberse sin ninguna duda si un corpus es fiable? Señalaba Le Roy Ladurie (1982 [1975]: 10) que “toute étude historique doit ou devrait commencer par

⁸ Sobre este tema resulta esclarecedor el manual de Pierazzo (2015) acerca de los diferentes acercamientos teóricos y metodológicos y modelos de edición digital de textos.