

# #23

# ¿QUÉ SIGNIFICACIÓN TIENEN LAS HUMANIDADES DIGITALES PARA LAS HUMANIDADES?\*

**Alan Liu**

*UC Santa Barbara*

**Traducido por Rocío Ortuño Casanova**

Ilustración || Ignacio Gómez Sales

Artículo || Invitado | Publicado: 07/2020

DOI: 10.1344/452f.2020.23.6

[ayliu@english.ucsb.edu](mailto:ayliu@english.ucsb.edu) / [rocio.ortuno@uantwerpen.be](mailto:rocio.ortuno@uantwerpen.be)

Licencia || Reconocimiento-No comercial-Sin obras derivadas 3.0 License



**Resumen** || El artículo busca demostrar que las humanidades digitales solo pueden ser explicadas y comprendidas si tenemos en cuenta que la problemática subyacente es la identidad no solo de las humanidades digitales, sino también de las humanidades en general. Tanto en lo que prometen como en la amenaza que suponen, las humanidades digitales sirven como simulacro de una manera futura de trabajar y estudiar las humanidades integrando algunos valores de la sociedad contemporánea sobre lo digital, sin perder su esencia en favor de otros dominios del conocimiento que se han vuelto digitales para reivindicarse en esa sociedad. El artículo explora mediante el análisis de un proyecto real, paso a paso, cómo la cuestión de la significación de las humanidades digitales refleja tanto un problema específico de las humanidades digitales como la crisis todavía mayor de la importancia y la significación de las humanidades de hoy en día.

**Palabras clave** || Humanidades | Tecnología digital | Humanidades digitales | Análisis digital de textos

**Abstract** || This article argues that the digital humanities can only rise to the level of an explanation if we see that the underlying issue is the disciplinary identity not of the digital humanities but of the humanities themselves. In both their promise and their threat, the digital humanities serve as a shadow play for a future form of the humanities that wishes to include what contemporary society values about the digital without losing its soul to other domains of knowledge work that have gone digital to stake their claim to that society. It explores by means of the analysis of a real project step by step how the question of the meaning of the digital humanities registers both a specific problem in the digital humanities and the larger crisis of the meaningfulness of today's humanities.

**Keywords** || Humanities | Digital technology | Digital humanities | Digital analysis of texts

**Resum** || L'article busca demostrar que les humanitats digitals només poden ser explicades i compreses si tenim en compte que la problemàtica subjacent és la identitat no sols de les humanitats digitals, sinó també de les humanitats en general. Tant en el que prometen com en l'amenaça que suposen, les humanitats digitals exerceixen de simulacre d'una manera futura de treballar i estudiar les humanitats que desitja incloure alguns valors de la societat contemporània sobre allò digital, sense perdre la seva essència en favor d'altres dominis del coneixement que s'han tornat digitals per a reivindicar-se en aquesta societat. Mitjançant l'anàlisi pas a pas d'un projecte real, l'article explora com la qüestió de la significació de les humanitats digitals reflecteix tant un problema específic de les humanitats digitals com la crisi encara major de la importància i la significació de les humanitats d'avui dia.

**Paraules clau** || Humanitats | Tecnologia digital | Humanitats digitals | Anàlisi digital de textos

Esta cuestión de la significación de la disciplina —que planteo desde el punto de vista de las humanidades en general— es más amplia que el problema de la identidad como disciplina que ahora preocupa tanto a las «HD», tal y como las llaman los concedores. Tras haber logrado una gran cantidad de acólitos, publicaciones, congresos, concursos para conseguir becas, tras haber llegado a multitud de instituciones (centros, programas y anuncios de puestos de trabajo), y haber sido ampliamente visibilizado, el campo está en pleno proceso de construcción de una identidad<sup>1</sup>. Los debates recientes sobre si las humanidades digitales son como una «gran carpa» (Jockers y Worthey, 2011), sobre «quién está dentro y quién se queda fuera» (Ramsay, 2008), si «tienes que saber extraer datos y crear herramientas para procesarlos [es decir, ser un constructor]» (Ramsay, 2011a), el debate sobre la necesidad de «más hackeo y menos rollo» [*more hack, less yack*] (Cecire, 2011b; Koh, 2012), y el de la presencia o ausencia de teoría en el campo expuesto en «¿a quién llamas tú ateórico?» [*who you calling untheoretical?*] (Bauer, 2011) atestiguan una dialéctica de inclusión y exclusión no tan diferente de la de otros campos emergentes en el pasado<sup>2</sup>. Un etnógrafo del campo, de hecho, podría inspirarse en alguna página de Claude Lévi-Strauss y representar las humanidades digitales actuales con una especie de diagrama de afinidades y diferencias entre tribus vecinas. Exagerando las diferencias un poco, como cuando una tribu se jacta de su singularidad, podemos decir que las humanidades digitales —que en muchos de sus aspectos se asocian con disciplinas de las viejas humanidades como la literatura, la historia, la filología clásica y las lenguas; con la recuperación de antiguos soportes como los libros o las bibliotecas; y en última instancia con el valor de lo antiguo en sí mismo (historia, archivos, las tareas de preservación)— no son la tribu de los «estudios de los nuevos medios de comunicación» [*new media studies*], bajo el influjo de las artes del diseño, visuales y mediáticas; la teoría continental; la crítica cultural; y la vanguardia de lo nuevo<sup>3</sup>. Del mismo modo, a pesar de las importantes tendencias que bogan hacia el trabajo en red y multimodal que abarca los medios sociales, visuales, auditivos y táctiles, gran parte de las humanidades digitales se centra en el estudio de documentos y de textos, de manera que el trabajo en este campo se distingue de la investigación digital que se hace en los estudios de los nuevos medios de comunicación, en las ciencias de la comunicación, en las ciencias de la información y en la sociología. Y las humanidades digitales están explorando nuevos repertorios de «crítica algorítmica» interpretativa o expresiva (la «segunda ola» de las humanidades digitales proclamada en «The Digital Humanities Manifesto 2.0», 3), de modo que incluso permite distinguir este campo de su precedente, las «humanidades computacionales» [*humanities computing*], del que se supone que tenía objetivos técnicos y de servicio más limitados y menos ambiciosos<sup>4</sup>. Recientemente, también se ha considerado un rasgo

## NOTAS

\* | Publicado y traducido con la autorización del autor. Artículo publicado originalmente en 2013 con el título «The Meaning of the Digital Humanities», en *PMLA*, 128, 2, 409-423, <<https://escholarship.org/uc/item/5gc857tw>>. **N. de la T.:** traduzco aquí la palabra *meaning* por «significación» como término que engloba «significado», «sentido» e «importancia» porque siendo el tema principal, aparecerá con todos estos significados a lo largo del artículo. Básicamente, se pregunta por la relevancia de las humanidades digitales y por los significados y conclusiones humanísticas que podemos extraer de ellas.

1 | Entre los trabajos panorámicos más recientes sobre las humanidades digitales se encuentran los artículos de Svensson (2009; 2010; 2011); Hayles (2012); Kirschenbaum (2010; 2012); Liu (2012b; 2012c); Burdick, Drucker, Lunenfeld, Presner, y Schnapp (2012); y «Digital Humanities Manifesto 2.0» (2009). Entre las colecciones de ensayos recientes o de próxima aparición sobre las humanidades digitales están Schreibman, Siemens y Unsworth (2004); Gold (2012); Price y Siemens (2013); y Goldberg y Svensson (2015). Sobre si las humanidades digitales son un campo y sobre el uso lingüístico de la frase «humanidades digitales», véase el artículo de Liu (2013b), «Is digital humanities a field?».

2 | Estas expresiones en conferencias recientes, blogs, tweets, hilos de discusión, documentos, etc., se han convertido en importantes «memes» que conforman la comunidad de las humanidades digitales. Sobre los temas de teoría, construcción, y *hack versus yack* (hackeo versus rollo),

distintivo la limitada conexión de las «humanidades digitales» con las cuestiones relacionadas con la identidad y con la justicia social, por parte, por ejemplo, del pujante colectivo #transformDH, al que le preocupa el hecho de que las humanidades digitales (a diferencia de algunas áreas de los *new media studies*) no se preocupen por cuestiones de raza, género, sexualidades alternativas o discapacidad<sup>5</sup>.

Evidentemente hay solapamiento de personas y métodos entre las tribus, hasta el punto de que, volviéndonos a inspirar en las páginas de Lévi-Strauss, habría que mencionar a ciertas figuras un poco tramposas que se sitúan entre dos de las tribus colindantes a las humanidades digitales y recogen características diferenciadoras de ambas que las distinguen de las humanidades digitales. Uno de esos, digámosles tramposillos, es el campo de la historia del libro. Prolífico en estudios que revisan textos clásicos, efímeros, marginales, la lectura social, la historia de la publicación, etc., la historia del libro ahora es en parte paralela a las ciencias de la comunicación pos-McLuhan. Peter Stallybrass afirma, por ejemplo, que «el código y el libro impreso eran los ordenadores de indexación que adoptó la cristiandad» así como su «tecnología disruptiva» o, como también se dice, su «innovación radical» (2002: 74,73). La historia del libro, por tanto, se solapa en la parte que tiene más similar a los estudios de los nuevos medios de comunicación con la arqueología de medios (especialmente en la tradición alemana de este campo, bien representada por el libro de Cornelia Vismann, *Files: Law and Media Technology*, 2008) y, del lado de las humanidades digitales, con aproximaciones a textos digitales basadas en la edición revisionista de textos, la bibliografía entendida como sociología de los textos y la materialidad de lo digital<sup>6</sup>. Otro de estos campos tramposillos es el de los estudios científico tecnológicos (ECT), que están curiosamente infrarrepresentados tanto dentro de los estudios de nuevos medios como en las humanidades digitales, incluso siendo claramente relevantes, como podemos inferir del trabajo de figuras como N. Katherine Hayles (2012) y Tim Lenoir. También la sociología histórica —que ahora aplica análisis de redes sociales a sus corpus de documentos— es uno de estos listillos que transita el camino medio entre las ciencias sociales y las humanidades digitales, incluso si estas últimas también están usando cada vez más análisis de redes sociales para estudiar obras teatrales, novelas, recepción literaria, historia etcétera<sup>7</sup>. Y de esos grupos de gente que alterna y entremezcla —como los tramposillos— las humanidades digitales y los estudios de nuevos medios, el grupo HASTAC —que ha tenido un éxito enorme aprovechando sobre todo las energías de los estudiantes de postgrado («académicos HASTAC»)— es quizás el más chaquetero.

## NOTAS

véase también la «Introduction» de Cecire (2011a).

3 | Dado que la cuestión de quién está incluido en las humanidades digitales puede ser un tema delicado, debo aclarar mi opinión sobre el alcance del campo. Aunque en este trabajo me aproximo a los estudios literarios digitales y me centro en un ejemplo concreto de investigación, las humanidades digitales son mucho más amplias. Por esta razón, a lo largo del trabajo voy a ir relacionando el ejemplo que tomo con métodos y problemas compartidos en todo el campo (incluida, por ejemplo, la historia digital). Sin embargo, aunque mi objetivo es abordar una noción bastante amplia de las humanidades digitales tal y como se practican y discuten habitualmente, no voy a tratar de ampliar tanto el círculo como para que entre todo lo que podría o debería quizás estar en él, pero que hasta ahora por lo general no se ha reconocido como parte de esta noción (por ejemplo, el campo del diseño), ya que ampliar demasiado el alcance podría acabar haciendo que el término humanidades digitales se convirtiera en algo informe o simplemente en una aspiración. Además, no trato específicamente el trabajo digital en ramas tan afines a las humanidades como son la arqueología y la antropología, que, a pesar de ser ramas de las ciencias sociales, se han unido a las humanidades en los llamados giros lingüísticos, giros culturales y tendencias relacionadas. El hecho de que ese trabajo se considere humanidades digitales depende de la cuestión previa de si, dejando de lado los métodos digitales, es humanístico (es decir, si encajan, por ejemplo, con la imprecisa definición de las

Incluso si completáramos nuestro hipotético diagrama etnográfico, este no conseguiría todavía explicar adecuada y completamente las humanidades digitales. Dejaríamos sin explicación la relación de las humanidades digitales con las humanidades en general. Mi tesis es que las humanidades digitales solo pueden ser explicadas y comprendidas si tenemos en cuenta que la problemática subyacente es la identidad no solo de las humanidades digitales, sino de las humanidades en general. Respecto a las humanidades, las humanidades digitales exceden (aunque incluyen) su papel funcional de instrumento o servicio, su papel pionero como fuente de innovación, el rol de «campo adicional» dentro del conjunto e, incluso, supera esos falsos roles políticos que se suelen asignar a los nuevos campos que a veces se ven como como desafiantes, renovadores, y (en un tono menos positivo) quintacolumnistas. Esto se debe a que las humanidades digitales también tienen un papel simbólico. Tanto en lo que prometen como en la amenaza que suponen, las humanidades digitales sirven como simulacro de una manera futura de trabajar y estudiar las humanidades integrando algunos valores de la sociedad contemporánea sobre lo digital, sin perder su esencia en favor de otros dominios del conocimiento que se han vuelto digitales para reivindicarse en esa sociedad. O, precisamente porque las humanidades digitales son tanto funcionales como simbólicas, una metáfora más acertada sería comparar las humanidades digitales con el registro de datos en una Unidad Central de Procesamiento (CPU) de un ordenador, en la que los valores almacenados en la memoria se cargan para poder mezclarlos, manipularlos y hacer pruebas con ellos muy rápidamente —en este caso, para probar las nuevas identidades de las disciplinas humanísticas que han evolucionado para ampliarse y abarcar más conocimientos y personas que trabajan con el conocimiento contemporáneo.

La cuestión de la significación de las humanidades digitales es la que mejor abre este debate porque refleja tanto un problema específico de las humanidades digitales como la crisis todavía mayor de la importancia y la significación de las humanidades de hoy en día.

La significación es claramente un metavalor y también un metaproblema para las humanidades digitales. Para desentrañar este problema de significación que tienen, destacaré un trabajo reciente de investigación literaria digital de dos académicos que están comenzando, que es de lo más novedoso y representativo de las principales tendencias en las humanidades digitales —una táctica que tiene la ventaja adicional de proporcionar a los que están

## NOTAS

regulaciones de la National Foundation on the Arts and Humanities de 1965: «aquellos aspectos de las ciencias sociales que tienen un contenido humanístico y emplean métodos humanísticos» (National Endowment, 1)]. Estas aclaraciones también se aplican a mis generalizaciones sobre los estudios de los nuevos medios de comunicación y a la burda simplificación que hago aquí (matizo la cuestión más adelante) para separarlos de las humanidades digitales. Agradezco al público de las charlas en las que presenté versiones anteriores de este trabajo que me ayudaran a pensar en estas cuestiones sobre la inclusión en las humanidades digitales, ya que algunos señalaron que no reconocían completamente su campo o trabajo en mi descripción. En última instancia la pregunta de «¿quién forma parte de las humanidades digitales?» se irá ajustando y afinando sobre el terreno a través de procesos normales profesionales de adjudicación: dónde se consigue un trabajo o se coloca a los estudiantes, en qué revistas y libros digitales o impresos se publica, en qué foros o blogs, a qué hashtags de Twitter se está asociado, a qué congresos se va, qué becas se reciben, etc.

4 | Sobre la crítica algorítmica, véase el ensayo de Ramsay (2008) con ese título.

5 | Sobre #transformDH, véase Phillips (2011). El colectivo ahora aparece a través de su hashtag en Twitter, tiene una página Tumblr, y es un grupo HASTAC.

6 | Ejemplos de la edición textual, la bibliografía como sociología de los textos, y la materialidad de los enfoques digitales en las humanidades digitales incluyen,

fuera de este campo un vistazo de principio a fin de un ejemplo de investigación realizada por humanistas digitales—. El trabajo es *A Quantitative Literary History of 2,958 Nineteenth Century British Novels: The Semantic Cohort Method* de Ryan Heuser y Long Le-Khac (2012b), el cuarto de los influyentes boletines digitales publicados por Laboratorio Literario de Stanford<sup>8</sup>. Heuser y Le-Khac informan sobre sus innovaciones en los métodos de «lectura distante» [«*distant reading*»] y minería textual [*text mining*] que son las señas de identidad del laboratorio (donde trabajaron con Matthew L. Jockers, Franco Moretti, y otros), y lo hacen con una autoconciencia metodológica que sitúa el problema de la significación por delante de lo demás. Afirman casi al principio:

¿Cuál es el *significado* de los cambios en la frecuencia del uso de palabras? ¿Qué hacemos con esos datos? Ahora que tantas investigaciones se basan en frecuencias de palabras y otros aspectos cuantificables de la cultura, estas son grandes preguntas. Ya estamos viendo que el mayor desafío a la hora de desarrollar métodos digitales de humanidades puede que no sea cómo extraer datos de objetos humanísticos, sino cómo analizar esos datos de manera *significativa* e interpretable (Heuser y Le-Khac, 2012b: 4; el énfasis es mío).

[[W]hat is the *meaning* of changes in word usage frequencies? What do we do with such data? With much current research drawing on word frequencies and other quantifiable aspects of culture, these are big questions. We can see now that the greatest challenge of developing digital humanities methods may not be how to cull data from humanistic objects, but how to analyze that data in *meaningfully* interpretable ways.]

Y añaden en sus conclusiones:

El problema metodológico general de las humanidades digitales puede ser planteado de manera muy simple y muy clara: ¿cómo pasamos de los números a *los significados*? Los objetos que se investigan, las pruebas que se recogen, las formas en que se analizan, todo eso es cuantitativo. Cómo pasar de este tipo de evidencias y de objetos a argumentos cualitativos y a reflexiones sobre temas humanísticos — cultura, literatura, arte, etcétera— no está claro (Heuser y Le-Khac, 2012b: 46; el énfasis es mío).

[The general methodological problem of the digital humanities can be bluntly stated: How do we get from numbers to meaning? The objects being tracked, the evidence collected, the ways they're analyzed—all of these are quantitative. How to move from this kind of evidence and object to qualitative arguments and insights about humanistic subjects— culture, literature, art, etc.—is not clear.]

A lo largo de todo el texto, Heuser y Le-Khac dan al *significado* una centralidad que indica que en las humanidades digitales el problema del significado tiene aproximadamente el mismo peso que el problema de «salvar los fenómenos» en la filosofía de la ciencia —es decir, el problema de relacionar los fenómenos observados empíricamente con fenómenos explicables, teorizables o predecibles

## NOTAS

respectivamente, McGann (2001); el TEI de la Iniciativa de Codificación de Textos, *TEI: P5 Guidelines*, con su atención a la prosopografía y las relaciones sociales («Nombres, fechas, personas y lugares»); y Kirschenbaum (2008).

7 | Entre los ejemplos de sociología histórica a los que me refiero se encuentran Franzosi (2010); Bearman y Stovel (2000); y Mohr y Duquenne (1997). Un ejemplo de análisis de redes sociales en los estudios literarios digitales es el de Moretti (2011). Un ejemplo de análisis de redes sociales en la historia digital es Lemercier y Rosental (2010). (Para otros historiadores que utilizan los métodos y herramientas del análisis de redes sociales, véase la «Bibliografía».)

8 | Véase también «Learning to Read Data» de Heuser y Le Khac (2011), que resume las investigaciones que figuran en su programa.

(«salvados»)—. En definitiva, Heuser y Le-Khac quieren «salvar los datos» haciéndolos significativos<sup>9</sup>. El tipo de datos que desean salvar es cuantitativo.

Evidentemente, ningún ejemplo puede tipificar todos y cada uno de los aspectos que forman parte del floreciente campo de las humanidades digitales. Por lo tanto, antes de examinar los pormenores del problema del significado en este caso, será útil hacer cuatro observaciones que relacionen la investigación de Heuser y Le-Khac con otros trabajos en las humanidades digitales. El problema de la significación puede no ser tan central en todas partes, pero su marco de análisis es convertible a otros marcos, de modo que podemos ver que muchas partes del campo conectan con el mismo tipo de cuestiones. A continuación se presenta una especie de tabla de conversión para relacionar el trabajo de Heuser y Le-Khac con otras áreas de las humanidades digitales.

En primer lugar, Heuser y Le-Khac seleccionan su material de investigación a partir de textos ya digitalizados (2958 novelas británicas escritas entre 1785 y 1895, todas extraídas de *The Internet Archive*, excepto 250 del repositorio Chadwyck-Healey)<sup>10</sup>. Esto significa que su trabajo no está dentro del área de la digitalización, la codificación de textos, la publicación o la archivística (actividades características de muchos proyectos de edición, recopilación y conservación digital), sino que entra dentro del área del procesamiento y análisis de repositorios digitales ya construidos (a la manera de los proyectos de análisis de textos, análisis de redes sociales, visualización, historia de los espacios, etc.). Sin embargo, las humanidades digitales son lo suficientemente jóvenes como para que estos dos amplios bloques de trabajo no estén totalmente especializados. Por una parte, los principales proyectos relacionados con la codificación de textos y el archivo digital suelen estimar necesario crear sus propias herramientas de análisis, procesamiento y visualización para presentar los materiales. Y, por otro lado, los proyectos relacionados con el análisis de texto, la visualización y otros tipos de procesado de datos a partir de textos, a menudo tienen que hacer un gran esfuerzo para seleccionar, limpiar y preparar materiales digitales preexistentes para lograr un corpus utilizable.

En segundo lugar, aunque Heuser y Le-Khac formulan su problema de significación en términos de cuantificación, esa no es la única formulación posible. La cuantificación es, en efecto, un problema clave en las humanidades digitales, especialmente cuando hablamos de esas cantidades ingentes que aparecen en los concursos para conseguir financiación de «*big data*» dentro del campo. Pero los humanistas digitales también tienen formas no cuantitativas de plantear el problema de la significación. Una de ellas es la idea de los modelos (y los procesos de modelado)

---

## NOTAS

9 | Para un análisis de la cuestión de «salvar los fenómenos» que tiene una importancia central en el tema de los datos, véase Bogen y Woodward (1988).

10 | La lista completa de las novelas del corpus de Heuser y Le Khac se puede encontrar en su «Online Companion» (2012a). La información sobre la fuente de los textos digitalizados proviene de un correo electrónico de Heuser al autor.

---

que ha sido ampliamente desarrollada por Willard McCarty (2005), entre otros. Los modelos revelan el significado (deconstruido en patrones, tendencias, formas), pero solamente a costa de reducir las dimensiones y las características del mismo. Los modelos diagramáticos, especialmente las visualizaciones que proliferan en las humanidades digitales (y que también vemos en el ensayo de Heuser y Le-Khac), son un ejemplo de ello, ya que las grandes redes de discurso (visualizadas a través de herramientas como *Gephi*) son comprensibles mientras su alcance (*scope*) o su nivel de detalle se mantienen bajos, pero, en cuanto uno de los dos parámetros se incrementa, se convierten en galaxias bellamente desconcertantes de nodos y enlaces. Otro ejemplo es el tipo de modelo textual que permite la codificación del texto (plasmado en definiciones de tipo de documento [DTD] y esquemas XML). Así pues, un debate bien conocido en las humanidades digitales se refiere a si el principio de una jerarquía ordenada de objetos de contenido (OHCO) que subyace a esos modelos hace que los textos sean legibles para la máquina únicamente a costa de rechazar toda la gama de lo que Jerome McGann denomina «superposición» [«*overlapping*»] y «estructuras recursivas» [«*recursive structures*»] legibles para el ser humano (McGann, 1999).

En tercer lugar, aunque nos centremos en la cuantificación y la tomemos como el mayor problema en torno a la significación, esta será diferente según la rama disciplinaria de las humanidades digitales que estemos tratando, ya que la historia digital y los estudios literarios digitales, por mencionar dos de las grandes áreas dentro del campo, llegan a la cuantificación por diferentes caminos y en diferentes etapas. La historia digital surgió en una disciplina que ya había experimentado con ímpetu el método numérico. Pensemos, por ejemplo, en la vertiente cuantitativa de la escuela *Annales* y en la cliometría. Tal vez como resultado de ello, la historia digital está más adelantada que la mayoría de las ramas de las humanidades digitales en la tarea de moldear el trabajo cuantitativo a una disciplina y una metodología que podemos considerar heredera del método de *Annales*: la historia cartográfica y de los espacios. Ahora mismo en la historia digital, por decirlo de alguna manera, lo más puntero es Fernand Braudel más la cartografía por satélite (Seefeldt y Thomas, 2009; Owens, 2007).

En cuarto lugar, la investigación de Heuser y Le-Khac requería una combinación de habilidades en programación e interpretación, que nos hace volver a la pregunta clásica que recientemente ha atormentado a las humanidades digitales: «¿hay que ser constructor?». Y las ha atormentado porque la respuesta, en efecto, es que hay que ser tanto «constructor» como «intérprete». Menos interesante que esta pregunta en sí misma, que se basa en una noción cada vez más obsoleta de trabajo en solitario en el que uno



---

es o bien constructor o bien intérprete, es la forma en que Heuser y Le-Khac son a la vez constructores e intérpretes, que es a través de una fructífera colaboración. Aquí llegamos a los límites exteriores del marco del problema de significado, donde se convierte en un marco coextensivo. Así como la significación es tanto un metavalor como un metaproblema, también lo es la colaboración, ya que tiene que ver con temas tan urgentes en las humanidades digitales como la coautoría, la construcción de proyectos colectivos, los libros multigráficos, la revisión abierta por pares, las redes sociales, el *crowdsourcing*, y las implicaciones que todos estos temas tienen en las cuestiones de contratación y de promoción. En lugar de tratar aquí en profundidad el problema de la colaboración, únicamente voy a destacar que es fundamentalmente convertible al problema de la significación. Por ejemplo, la cuestión de qué tipo de conocimiento es producido por «la sabiduría de las masas», «la inteligencia colectiva», «la larga cola», «la conciencia de la colmena», «la folksonomía», etc. (todos ellos memes dominantes de la Web 2.0) es esencialmente una cuestión sobre la significación, la importancia de la versión social de los grandes datos, la gran multitud. La conciencia, o la inconsciencia, de esa multitud ha sido un problema central de la modernidad por lo menos desde la Revolución Francesa.

Usando estas observaciones como guía para relacionar el trabajo de Heuser y Le-Khac con otros enfoques de las humanidades digitales, podemos ahora mirar de forma más exhaustiva su boletín. El problema de la significación se pone de manifiesto en este boletín como consecuencia de una de las innovaciones clave de Heuser y Le-Khac: el uso de *Correlator*, una herramienta que han creado con la que a partir de «palabras-semilla», sugeridas en parte por «la erudición literaria existente» (2012b: 11), han procesado su extenso corpus de novelas del siglo XIX (con la ayuda de una base de datos diseñada por Jockers que tabulaba el número de ocurrencias de cada palabra [Heuser y Le-Khac, 2012b: 6]) para encontrar otras palabras que estuvieran estadísticamente correlacionadas con las palabras-semilla y cuyas tendencias de frecuencia, medidas diacrónicamente a lo largo del siglo en intervalos de décadas, siguieran de cerca las tendencias de frecuencia de las palabras-semilla. En otras palabras, *Correlator* encuentra lo que Heuser y Le-Khac llaman cohortes o «conjuntos de palabras» en el corpus, que consisten en palabras que suelen ir juntas y que se comportaron de manera similar a lo largo del tiempo, aumentando o disminuyendo su frecuencia juntas y destacándose de otros conjuntos de palabras. El resultado final —al que llego saltándome los detalles del método algorítmico de Heuser y Le-Khac, así como, de momento, un importante paso de ajuste en su método— es la identificación en los conjuntos de palabras de «campos semánticos ricos y consistentes» que son «tanto semántica como culturalmente legibles» en relación con las tendencias históricas. Un ejemplo son las 136 palabras como *gentil*,

*sensato, vanidad, elegante, delicadeza, reserva, suave y moderación* que conforman el «campo de la moderación social» en las novelas (2012b: 8).

Usando este método, Heuser y Le-Khac hicieron dos descubrimientos principales. En primer lugar, utilizaron las «palabras semilla» *integridad, modestia, sensibilidad y razón* (convenientemente adecuadas para la investigación sobre novelas de esa época) para encontrar un conjunto sorprendentemente grande de «palabras abstractas, socialmente normativas, evaluativas y altamente polarizadas», cuyas frecuencias disminuían drásticamente conforme avanzaba el siglo. Este conjunto se subcategorizó en campos semánticos etiquetados como «restricción social», «valoración moral», «sentimiento» y «parcialidad» (2012b: 11-19). Luego, por casualidad (obviando las investigaciones previas que se habían hecho sobre el tema), introdujeron la improbable «palabra semilla» *duro [hard]* en el *Correlator* para descubrir un conjunto sorprendentemente distinto, muy grande, de «palabras descriptivas, muy concretas, de uso directo y cotidiano», cuyas frecuencias aumentaban considerablemente con el tiempo. Este conjunto lo subcategorizaron en campos semánticos para verbos de acción, partes del cuerpo, colores, números, adjetivos y preposiciones de localización y dirección, y adjetivos físicos (2012b: 19-27). La correlación inversa tan fuerte que había entre estos dos grandes conjuntos diferenciados de palabras, que además corroboraron estadísticamente (2012b: 28-29), les dio las pistas necesarias para empezar a interpretar. Concluyeron que el conjunto de palabras «abstractas» constaba de palabras cuyo uso, si bien no estaba anclado en lo específico, estaba «controlado y estrechamente limitado» a las comunidades rurales tradicionales más pequeñas, representadas más o menos a principios de siglo en las novelas evangélicas, góticas y de aldeas y en las novelas de Jane Austen, Walter Scott, y George Eliot (Eliot es un atípico cronológico); mientras que el conjunto de palabras relacionadas con «duro» [*hard*] está poblada por palabras cuya referencialidad y alienación de contextos más amplios se correlacionaron más adelante en el siglo con los «espacios sociales más amplios y menos restringidos» de los centros urbanos representados en las novelas urbanas, industriales, de aventuras, de fantasía, de ciencia ficción y para niños (2012b: 30-34).

Esta idea finalmente llevó a Heuser y Le-Khac al nivel más alto de interpretación: siguiendo las obras *Cultura y sociedad* y *El campo y la ciudad* [*Culture and Society* y *The Country and the City*] de Raymond Williams, sugirieron que las tendencias inversas del tipo A («abstractas») y del tipo B (las relacionadas con la palabra «duro») revelaban algo significativo sobre el «espacio social de la novela». De este modo, afirman que los «valores de conducta y normas sociales» en las «comunidades conocibles» (una frase

de Williams), expresadas dentro del conjunto del tipo A declinaban ante «la urbanización, la industrialización y las nuevas etapas del capitalismo», expresadas en el conjunto del tipo B (2012b: 33, 35-36). La verificación computacional de esta tesis previamente conocida, junto con el descubrimiento de conjuntos de palabras precisas que dan una visión genuinamente nueva de la tesis, les permitió ofrecer una crítica literaria y cultural más normativa y más reconocible en secciones posteriores de su boletín, en el que se refieren a la acción, el entorno y los personajes (2012b: 37-45). Aquí leen atentamente los textos (e incluyen, por primera vez, citas de los mismos), ajustan las tendencias agregadas a «unidades comprensibles y familiares para nosotros como lectores y estudiosos de la literatura[:] las novelas, los géneros y los autores reales» (2012b: 31), y generalizan sobre los cambios radicales en la historia cultural, pero con la diferencia fundamental de que su lectura no se basa en las habituales observaciones anecdóticas, el falso empiricismo o las asunciones a partir de un caso único habituales en la crítica literaria (me refiero por ejemplo a cuando los críticos afirman que una palabra aparece «a menudo» o en un lugar «importante» sin haberlo comprobado cuantitativamente), sino que basan sus líneas de interpretación a partir de la observación de los resultados generados por el análisis computacional.

Aquí es donde llegamos al quid de la cuestión de la significación en las humanidades digitales. Mientras Heuser y Le-Khac mezclan «procedimientos supervisados y no supervisados» (2012b: 28), la parte más nueva, audaz e interesante de su metodología es la no supervisada. Esta contribuye al desarrollo de un objetivo importante y general de las humanidades digitales que podría llamarse «interpretación partiendo de tabula rasa», es decir, que la interpretación se lleva a cabo exclusivamente a través de lo que se va descubriendo, sin tener en cuenta las hipótesis previas acerca los fenómenos. Este método también es compartido por otros métodos de las humanidades digitales como el *topic modeling* (un método estadístico para agrupar palabras en torno a diferentes temas que Heuser y Le-Khac utilizan para corroborar sus hallazgos). Lo más puramente ideal sería alcanzar lo que Heuser y Le-Khac llaman «un método no supervisado que genera temas sin la intervención subjetiva de los usuarios» (2012b: 28)<sup>11</sup>. Es decir, que un ordenador debe ser capaz de leer textos algorítmicamente y descubrir conjuntos de palabras que lleven a pensar en temas sin actuar sobre un concepto inicial de un intérprete que busque confirmar un tema en particular. Por supuesto, Heuser y Le-Khac suponen que hay temas preexistentes que se encuentran en los conjuntos de palabras de los textos primarios y que la misión principal es descubrirlos. Sin embargo, la interpretación partiendo de tabula rasa es igualmente un objetivo de la vertiente más posmoderna de las humanidades digitales, que defiende que los críticos deberían utilizar métodos

---

## NOTAS

11 | Goldstone y Underwood (2012) ofrecen una explicación y ejemplo de modelado de temas de especial interés para los lectores de *PMLA*. Por otro lado, Underwood explica más detalladamente la metodología. Para una discusión sucinta sobre la diferencia entre la minería de datos supervisada y la no supervisada —una distinción que se origina en el campo del *machine learning*— véase el «Analytics».

algorítmicos para jugar con los textos de forma experimental, generativa o «deformativa», de manera que se puedan descubrir formas alternativas de significado que no sean tan propias de las señales preexistentes como fieles a los riffs, a las tonadas repetitivas, a las líneas subyacentes a esas señales. El objetivo común es desterrar, o al menos reducir y aplazar bastante, las ideas preconcebidas en los inicios de la interpretación.

Sin embargo, la interpretación partiendo de tabula rasa pone en duda el objetivo final de Heuser y Le-Khac, que es pasar de los números a un significado humanístico («argumentos cualitativos e ideas sobre temas humanísticos —cultura, literatura, arte, etc.»). No está claro epistemológica, cognitiva o socialmente cómo los seres humanos pueden tomar una señal descubierta por una máquina y desarrollar una interpretación que conduzca a un concepto humanamente comprensible, a menos que esa señal (para ser reconocida como tal) contenga un origen conceptual coevaluado que sea conocible desde un principio porque, como mínimo, el intérprete humano ha conocido su forma o su posición (el hueco o lugar aproximado en el sistema semántico donde se espera que se aclare su significado, o al menos su pertenencia al sistema)<sup>12</sup>. Dicho de otra manera: si la máquina puede descubrir conjuntos de palabras surgidos a partir de palabras-semilla, ¿qué conceptos-semilla —o sea, semántica-semilla (utilizando «semántica», por el momento, como superpuesta a «conceptos» de manera coherente con el uso de Heuser y Le-Khac)— subyacen en el texto escondidos como forma latente y clasificatoria de posiciones semánticas relacionales capaces de convertir conjuntos de palabras en «campos protosemánticos» (2012b: 7)?

De ahí la inmensa importancia del paso de ajuste en el método de Heuser y Le-Khac que antes he eludido. De hecho, Heuser y Le-Khac usaron *Correlator* para producir solo conjuntos preliminares de palabras y no campos semánticos acabados porque se dieron cuenta de que antes debían asegurarse de que sus conjuntos tuvieran una consistencia semántica que la correlación cuantitativa por sí sola no podía ofrecer. Algunos de los conjuntos de palabras descubiertos por la máquina, por ejemplo, sugerían solo campos semánticos difusos que parecían quedar poco definidos por culpa de palabras extrañas que se incluían en él o, por el contrario, porque les faltaban palabras que deberían haber estado ahí. Dicho de otra manera, para que tuvieran sentido, los conjuntos de palabras tenían que ser filtrados y completados. De este modo, Heuser y Le-Khac se dieron cuenta de que no solo necesitaban «palabras semilla» sino «conceptos semánticos semilla», que precisamente quedaban suspendidos en lo que antes he llamado una «forma clasificatoria de posiciones semánticas relacionales», de manera que dotaban a los conjuntos de una idea o un sentido de lo que está más o menos próximo de un significado dado.

---

## NOTAS

12 | Sculley y Pasanek (2008) examinan el problema de la «circularidad» (y otras cuestiones) en la extracción de datos de las humanidades.

En resumen, Heuser y Le-Khac necesitaban un tesoro, y no cualquier tesoro, sino uno que ofreciera una semántica histórica que coincidiera con la dimensión longitudinal de sus conjuntos de palabras. Cuando ya habían empezado a utilizar el *Correlator*, se sacaron un as de la manga, un conejo de la chistera. Recurrieron al fantástico *Historical Thesaurus* del *Oxford English Dictionary* (2009; en adelante *HTOED*), que acababa de publicarse, y tomaron prestadas sus clasificaciones semánticas históricas a través de lo que ellos denominan «un método dialógico que se basaba tanto en datos históricos cuantitativos como en rúbricas semánticas cualitativas para construir campos semánticos con precisión y con las sutilidades necesarias»:

Habiendo pasado por una etapa de desarrollo del campo semántico durante la que habíamos adoptado un enfoque más centrado en lo empírico y en lo histórico, tuvimos que volver al enfoque semántico para hacer que esos conjuntos de palabras puramente empíricos fueran interpretables y significativos. Nuestro planteamiento inicial fue hacer un filtrado de estas palabras en busca de grupos que parecieran semánticamente coherentes, pero resultó ser demasiado impreciso y subjetivo [...]. Finalmente nos dirigimos al *[HT]OED* [...]. Es de lo más exhaustivo, sus categorías son sutiles y específicas, y está realmente organizado en torno al significado. Utilizamos esta potente taxonomía para hacer dos cosas: primero, ser más específicos en la identificación de las categorías semánticas que constituían nuestros conjuntos de palabras; segundo, ampliar estos conjuntos de palabras con muchas más palabras (2012b: 7)<sup>13</sup>.

[Having moved through an empirically and historically focused stage of semantic field development, we needed to return to the semantic focus in order to make such purely empirical word cohorts interpretable and meaningful. Our initial approach was to filter through these words for groups that seemed semantically coherent, but this proved too loose and subjective [...]. Finally we turned to the *[HT]OED* [...]. It's nearly exhaustive, its categories are nuanced and specific, and it's truly organized around meaning. We used this powerful taxonomy to do two things: first, be more specific in identifying the semantic categories that constituted our word cohorts; second, to expand these word cohorts with many more words].

Creado en la Universidad de Glasgow en 1965, el *HTOED* clasifica la lengua inglesa bajo tres conceptos semánticos principales, «el mundo exterior», «la mente» y «la sociedad», y luego, por etapas, el árbol clasificatorio va descendiendo, dividiéndose en miles de ramificaciones. En las ramas inferiores de la taxonomía, los sinónimos aparecen en el orden cronológico en el que entraron en el idioma<sup>14</sup>. Por lo tanto, a primera vista, la *HTOED* es la guía perfecta para buscar conceptos con los que alimentar la máquina de caza de palabras de Heuser y Le-Khac. Esto se hace aún más evidente cuando nos damos cuenta de lo poderosa que es la interpretación semántica involucrada en la creación del *HTOED*. Como se documenta en los informes de dos de sus jefes de redacción, Christian Kay e Irene

## NOTAS

13 | Para una descripción más completa de la forma en que Heuser y Le-Khac utilizaron el *HTOED*, véase el apéndice C de su *Quantitative Literary History* (2012b).

14 | Para este ensayo, he consultado el *HTOED* en línea, que ahora aparece como parte del OED en línea. Sobre el sistema de clasificación de la *HTOED*, véase «Structure of the Historical Thesaurus» (2009). Para información general adicional sobre la obra, véase Kay, «What Is the *Historical Thesaurus?*» (2012).

Wotherspoon, y en un detallado correo electrónico que me envió el actual director asociado del *HTOED*, Marc Alexander, la clasificación de las palabras por su significado y cronología por parte de los editores recordaba a los famosos papelitos que se usaron en la confección del propio *Oxford English Dictionary (OED)*<sup>15</sup>. El trabajo empezó antes de que la informática tuviera realmente un uso práctico para las humanidades, así que para hacer el *HTOED* unos cuantos seres humanos tuvieron que escribir durante décadas cada palabra del *OED* en hojas de papel con significados, fechas de uso y metadatos dispersos, para luego clasificar, agrupar y archivar las hojitas en grupos y jerarquías conceptuales. Cuando los ordenadores entraron en escena, lo hicieron originalmente con un papel secundario (para impulsar la tirada impresa de la obra). En su estado formativo, el *HTOED* se basó en una labor humana de ordenación semántica.

Al instalar el *HTOED* como si fuera un *plug-in* para el *Correlator*, Heuser y Le-Khac sembraron su proceso hermenéutico con una pizca de interpretación semántica humana. De este modo, «resolvieron» el problema de la significación hábilmente, al hacer de la aporía entre la interpretación cuantitativa a partir de tabula rasa y la interpretación cualitativa humanamente significativa su propia solución: un «enfoque dialógico que oscila entre lo histórico y lo semántico, entre las frecuencias empíricas de las palabras que revelan sus tendencias históricas y las taxonomías semánticas que nos ayudan a identificar el significado y el contenido de esas tendencias» (2012b: 9). Y todavía añaden: «En realidad, los métodos aquí desarrollados no se centran en las cohortes de palabras, que a pesar de que poseen una coherencia histórica pueden carecer de coherencia semántica, o en los campos semánticos que, aunque guarden una coherencia semántica, no observan conexiones históricas entre ellas. El verdadero objeto de estudio es un híbrido que satisfaga ambos requisitos» (2012b: 9-10). Esa hibridación es una característica predominante del método digital de las humanidades<sup>16</sup>.

¿Cómo podría este método híbrido estar mejor fundamentado teóricamente? Vuelvo a mi anterior teoría de la diferenciación de las humanidades digitales respecto a otros campos colindantes. Para elaborar una respuesta son relevantes dos de las carencias de las humanidades digitales reveladas por el diagrama etnográfico imaginario de campos que he trazado antes. La primera es la teoría y práctica del diseño, que Anne Burdick, Johanna Drucker, Peter Lunenfeld, Todd Presner y Jeffrey Schnapp, en su reciente *Digital\_Humanities*, declaran con la audacia propia de un manifiesto que es sumamente necesaria para el campo (2012: esp. 12-16, 117-119). Hoy en día, tengo la sensación de que la unión entre las humanidades digitales y los grupos de estudios de nuevos medios de comunicación que se necesitaría para la plena realización de esta visión es más

---

## NOTAS

15 | Para información sobre la elaboración del *HTOED* a la que se hace referencia más adelante, he consultado el artículo de Kay (2010 y 2012), «Classification» y «What Is the *Historical Thesaurus?*»; Kay y Chase (1987); y Wotherspoon (1992 y 2010); así como el correo electrónico de Alexander. Para ver las fotos de las hojas de papel, los cajones de archivo y los ordenadores utilizados, véase la *Photo Gallery* del *Historical Thesaurus*. Mi especial agradecimiento a Alexander por ayudarme a reunir recursos y por su correo electrónico lleno de detalles e historias sobre los procesos manuales y tecnológicos utilizados en la fabricación del *HTOED*.

16 | Véase, por ejemplo, la discusión de Gibbs y Cohen sobre su método híbrido (2011: 70, 76).

---

un deseo que una realidad, que existe solamente en proyectos específicos y no de manera programática. Sin embargo, los coautores del libro, que expresan una visión de las humanidades digitales muy propia de la Costa Este de Estados Unidos y las presentan como un campo muy arraigado en los medios de comunicación y las artes del diseño, presentan acertadamente el diseño como un principio de descubrimiento y generación de conocimiento en lugar de (como suele hacerse en las humanidades digitales) como una representación a posteriori de datos en diagramas de dispersión, gráficos de redes sociales y otras visualizaciones anticuadas, o en interfaces de publicación como libros o blogs que en realidad están también muy manidas. Como dicen los coautores, cuando «se utiliza para plantear y enmarcar preguntas sobre el conocimiento», el diseño es «un método intelectual», una «encarnación del argumento y la metodología de un proyecto», «un acto de pensamiento» y un «nuevo fundamento para la conceptualización y producción de conocimiento» (2012: 13, 14, 15, 117). El diseño interactivo, multimodal, dinámico y participativo en la era digital es un método no solo de reconocimiento de patrones, sino de *comprensión* de patrones. Ver el diseño en los datos es una manera de conocer el significado en las humanidades digitales.

La otra deficiencia a la que me refiero es la de los estudios científico-tecnológicos (ECT), que los humanistas digitales a menudo omiten incluso cuando hablan de «tecnología digital», «tecnología de medios», etc., como si la tecnología fuera una parte indivisible de lo digital y de los medios y no tuviera su propia historia, filosofía, sociología, política, economía y estética, todas entremezcladas con la ciencia, pero también distintas de ella. Me refiero sobre todo a la rama postmoderna de los ECT (por ejemplo la representada por Feyerabend, 1978; Latour, 2005; Pickering, 1995), cuyo posicionamiento «contra el método» de la ciencia (especialmente en sus extrañas relaciones con la tecnología) se basa en que cualquier búsqueda de un método estable para comprender cómo se genera el conocimiento por parte de los seres humanos utilizando máquinas se construye sobre la falacia inicial de que existen órdenes humanos y maquínicos inmaculadamente separados, cada uno con una pureza ontológica, epistemológica y pragmática que permite llevarlos a una relación metodológica cognoscible con el otro, ya sea una relación de amo y esclavo, causa y efecto, agente e instrumento, u otra. ¿Qué podríamos aprender de los ECT si tomáramos al pie de la letra lo que dice el Laboratorio Literario de Stanford y otros centros y programas digitales de humanidades y los estudiáramos como laboratorios, de la misma manera que Andrew Pickering estudió la «caza del quark» en un laboratorio de física (1995: 68-112)? La respuesta es que probablemente el método de las humanidades digitales —que converge, pero también a veces diverge del método científico— consista en coajustar repetidamente los conceptos

humanos y la tecnología de las máquinas hasta que (como en la tesis de Pickering sobre «la desfiguración de la práctica» [«*the mangle of practice*»]) ambos se establezcan mutuamente en posturas temporales sobre la verdad que ninguno de los dos podría sostener por sí mismo. El conocimiento es la danza de un patinador sobre una superficie epistémica resbaladiza, sobre la que no pueden estar solos ni el humano ni la máquina —ni el bailarín ni los patines—. Los ECT, en definitiva, son otro método para conocer el significado en las humanidades digitales. De hecho, se puede pensar que complementa el método del diseño. Después de todo, la diferencia entre la desfiguración de la práctica y un baile es el diseño.

De hecho, una aproximación desde los ECT podría abrir un capítulo fascinante en nuestra lectura de la obra de Heuser y Le-Khac que solo voy a poder relatar de forma muy breve<sup>17</sup>. Aunque no he realizado una investigación etnográfica y documental al estilo de los ECT en el Laboratorio Literario de Stanford (excepto a través de observaciones informales durante una visita de un día por invitación del laboratorio)<sup>18</sup>, he observado el «laboratorio» de la fabricación del *HTOED* desde la distancia, es decir, el laboratorio de semántica humana que Heuser y Le-Khac incorporan como socios de su laboratorio de análisis computacional. Cuando vemos el *HTOED* de esta manera, nos damos cuenta de que Heuser y Le-Khac no sacaron cualquier conejo de la chistera, sino que sacaron un conejo muy especial y muy parecido al que acabó haciendo caer a Alicia por la madriguera. El *HTOED* no supone tanto una solución al problema del significado como una versión recurrente, de espejo del mismo problema. Al menos cinco aspectos de su fabricación son relevantes en este sentido:

1. Aunque la creación del *HTOED* fue esencialmente precomputacional, no fue pretecnológica. Al requerir que los seres humanos escribieran palabras, significados y metadatos en hojas de papel y que luego clasificaran las hojas en cajones (y más tarde en bases de datos), el *HTOED* se originó como un acto semántico que entretecía totalmente lo tecnológico y lo humano, en lugar de ser simplemente un acto humano.

2. Siguiendo a Heuser y Le-Khac, hasta ahora he tratado la «semántica» y los «conceptos» como nociones coincidentes. Pero al referirse al método del *HTOED*, Kay separa la semántica léxica (relaciones de significado entre las palabras) y la semántica conceptual (significados vinculados a referentes del mundo exterior). Centrarse en la primera, dice, fue «esencialmente el procedimiento adoptado en el *HTOED*, y es a lo que nos referimos al decir que la clasificación debería “emerger” de los datos», es decir, los datos puramente léxicos dentro de su texto de origen, el OED (Kay, 2010: 265-266).

---

## NOTAS

17 | Resumo aquí una explicación más extensa del *HTOED* que incluiré en una versión de este trabajo que aparecerá en el libro que estoy escribiendo sobre humanidades digitales. Mi agradecimiento a Clare Birchall, quien en una conversación conmigo después de presentar una versión anterior de este trabajo en el King's College, hizo una pregunta que me llevó a empezar a investigar las dimensiones tecnológicas de la *HTOED*.

18 | Mi agradecimiento al Laboratorio Literario de Stanford por invitarme a visitarlo el 21 de mayo de 2012. Sobre la génesis y los principios del laboratorio, véase Jockers (2013).



3. La frase «“emerger” de los datos» es la clave de un importante principio del OED. Como afirma Kay, «nuestra posición teórica sobre el *HTOED* siempre ha sido que la clasificación en cualquier nivel debe desarrollarse a partir de los datos en lugar de ser impuesta siguiendo algún esquema predeterminado» (Kay, 2010: 258). Resulta que la autoridad subyacente del *HTOED* son los «datos» (el registro del idioma inglés observado a través del instrumento del *OED*). Es decir, el propio *HTOED* se adhirió al principio de descubrimiento de fenómenos a partir de tabula rasa, y de inicio a partir de ahí de la interpretación (en este caso, de interpretación taxonómica).

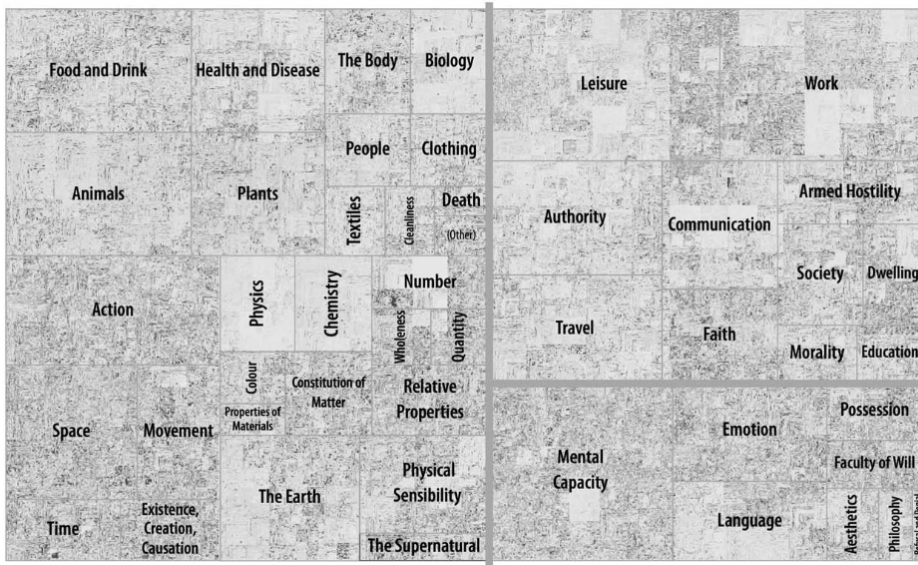
4. Así pues, después de estudiar los datos léxicos, los editores del *HTOED* contravinieron el esquema de ordenamiento de su predecesor canónico, el *Tesaurus de Roget* [*Roget's Thesaurus*], que había priorizado las «relaciones abstractas». El *HTOED*, sin embargo, prioriza el «mundo exterior» en su estructura taxonómica trinaría. Eso se debe a que los sentidos concretos y cercanos a la palabra que se relacionan con el mundo exterior entraron antes en el lenguaje (Kay, 2010: 257-262). Las implicaciones para nuestra lectura de Heuser y Le-Khac son sorprendentes. La amplia tendencia histórica identificada por el *HTOED*, según la cual las palabras concretas preceden a las abstractas, es el reverso de la tendencia novelística del siglo XIX que identifican Heuser y Le-Khac, en la que las palabras «abstractas» son las dominantes en una primera etapa y las «duras» después. Por supuesto, esto no contradice en modo alguno su tesis, ya que, en una enunciación más completa, su explicación podría ser que, debido a la urbanización, el siglo XIX (y la forma novelística) fue excepcional, o al menos solo coincidió con tendencias que se habían dado en muy pocos siglos en toda la historia (y la forma de la novela coincidió con solo unas pocas formas). Pero el cometido de leer las novelas del siglo XIX de forma conjunta con el corpus del *HTOED*, que es más amplio, requeriría realizar más comprobaciones para ver cuán excepcional era realmente la tendencia del siglo XIX y, por lo tanto, si deben tenerse en cuenta otros fenómenos socioculturales o de otro tipo para explicar su singularidad<sup>19</sup>.

5. Por último, casi como apéndice, debo añadir que, si bien la computación fue una idea tardía en el proceso de confección del *HTOED*, hoy en día es crucial para los usos avanzados de la obra. Los editores del *HTOED* migraron en algún momento su contenido a una serie de bases de datos relacionales porque se dieron cuenta de que la computación podría permitir consultas avanzadas en tiempo real (Wotherspoon, 2002: 218). De forma todavía más espectacular, un reciente ensayo de Alexander muestra que consultar el *HTOED* para conocer la historia de la lengua puede ser en sí mismo una forma de practicar humanidades digitales (Alexander, 2010). Alexander lleva a cabo un análisis computacional cuantitativo del *HTOED* para

## NOTAS

19 | En su análisis estadístico de la dicción en las obras literarias inglesas aparecidas a lo largo de un período algo más extenso (los siglos XVIII y XIX), Underwood y Sellers (2012) hacen un apunte relevante relacionado con esto, al comparar sus hallazgos con los de Heuser y Le-Khac. Realizan un seguimiento del aumento en el siglo XIX de la incidencia proporcional de las palabras de la época anterior, a menudo anglosajonas, en los géneros literarios (incluida la ficción en prosa) que es «en gran medida consustancial» con la creciente incidencia de palabras «duras» en las novelas de ese siglo encontradas por Heuser y Le-Khac. Sin embargo, agregan que la mayor longitud de la línea histórica que supone su punto de partida muestra que la trayectoria de tales palabras «había invertido recientemente su tendencia» como parte de una nueva forma de ser «literaria». «La relativa escasez de verbos de acción simples en la escritura de principios del siglo XIX, por ejemplo», señalan, «fue un acontecimiento reciente» en el lenguaje literario que revela transformaciones sociales solo a través de la mediación de «ideales rivales de refinamiento literario». Mi agradecimiento a Le-Khac por su e-mail aconsejándome el artículo de Underwood y Sellers, que no conocía.

generar visualizaciones que conducen a hipótesis sobre algunos cambios que han tenido lugar en el lenguaje (véase la imagen de abajo). Al conectar los ECT y las aproximaciones al diseño que he esbozado anteriormente, podríamos incluso decir que en este punto el uso del *HTOED* se convierte en un experimento de diseño digital. Al final, el *HTOED* no es eso «otro» que Heuser y Le-Khac necesitan para contribuir a que el trabajo de su *Correlator* sea significativo, sino que es en sí mismo el precursor del *Correlator*.



1. Visualización de Marc Alexander del inglés contemporáneo en el HTOED. Cada punto representa una palabra, y su sombra corresponde al momento de entrada de la palabra en el idioma (los puntos más oscuros muestran palabras más antiguas). Las palabras están ordenadas por proximidad semántica como se indica en las etiquetas.

Aunque bajar por la madriguera del conejo de Alicia es necesario para rastrear el fundamento metodológico de cualquier investigación llevada a cabo por las humanidades digitales, no es la forma adecuada de concluir este artículo. Así que me salgo del ámbito específico de mi ejemplo, e incluso del campo de las humanidades digitales, para abrir mi argumento a una aplicación más genérica con la siguiente reflexión de N. Katherine Hayles:

Cuanto más se avanza a lo largo del espectro que se dirige hacia la «*machine reading*» [lectura computacional], más se acepta implícitamente la creencia de que los eventos multicausales a gran escala son causados por confluencias que incluyen multitud de fuerzas que interactúan simultáneamente, muchas de las cuales no son humanas [...]. Es cierto que los eventos ocurren en una magnitud que excede en mucho a los actores individuales y que sobrepasa en mucho la capacidad de los humanos de asimilar la información relevante, sin embargo, la «*machine reading*» podría ser un primer paso hacia la creación de patrones visibles que la lectura humana podría interpretar posteriormente (2012: 29).

---

[The further one goes along the spectrum that ends with “machine reading,” the more one implicitly accepts the belief that large-scale multicausal events are caused by confluences that include a multitude of forces interacting simultaneously, many of which are non human [...]. If events occur at a magnitude far exceeding individual actors and far surpassing the ability of humans to absorb the relevant information, however, «machine reading» might be a first pass toward making visible patterns that human reading could then interpret (2012: 29)].

Ahora puedo revelar que no es casualidad que al principio de este ensayo aludiera a Lévi-Strauss y a la antropología estructural. El estructuralismo es un punto intermedio en el largo camino moderno hacia la comprensión del mundo como sistema (por ejemplo, como modos de producción; burocracia weberiana; lenguaje saussureano; sociedad de masas, medios de comunicación y sociedad corporativa; neoliberalismo; y así, sucesivamente) que ha obligado a que la vertiente progresista de las humanidades se separe de las humanidades anteriores, las del espíritu humano (*Geist*) y del yo humano, para adoptar una visión del mundo en la que, como dice Hayles (2012), «los eventos multicausales a gran escala son causados por confluencias que incluyen una multitud de fuerzas [...] muchas de las cuales no son humanas». Este es el telón de fondo en el que podemos ver cómo el problema de la significación de las humanidades digitales recoge la crisis general que hay hoy en día en torno a la relevancia de las humanidades. La crisis general se basa en que el significado humanístico, con sus anhelos residuales en torno al espíritu, a la humanidad y el yo —o, como decimos ahora, la identidad y la subjetividad— debe competir en el sistema mundial con los conocimientos sociales, económicos, de ingeniería científica, de trabajo y de cultura popular que no necesariamente valoran el significado o, dicho de forma que todavía da un poquito más de miedo: valoran el significado pero lo enmarcan sistémicamente en formas que alienan o cooptan el significado humanístico. Por lo tanto, el conocimiento humanístico actual se asimila cada vez más a lo que los propios humanistas llaman investigación, pruebas, análisis, método, productividad e «impacto» (ya que este término está institucionalizado en los «ejercicios de evaluación de la investigación» o «Research Assessment Exercises» de las universidades británicas), sin que quede tiempo ni espacio libre para ningún viejo fantasma en la máquina —a menos que, como he argumentado en *Laws of Cool* (2004), siga habiendo un anhelo de ser *cool*, ojalá que entre nuestros estudiantes, atrapados en la maraña educativa. La gente *cool* diga NO al sistema actual de conocimiento y trabajo, incluso mientras se meten en el cubículo que les corresponde entre las otras decenas de cubículos iguales de sus oficinas.

---

Claro que, si se tratara solo de un problema de metodología de la investigación, sería un poco exagerado llamarlo «crisis» de la significación y trascendencia de las humanidades. Sin embargo, la palabra «crisis» se vuelve apropiada cuando nos damos cuenta de que el problema de la significación también afecta a la pedagogía y a los empleos a raíz de la recesión económica, lo que hace que el problema afecte de manera cruel a los humanistas en su proceso formación o en la búsqueda de empleo (por no hablar de cómo afecta a los programas de humanidades que forman y emplean a estos humanistas). Aquí se abre un conjunto de cuestiones sobre las que no me puedo detener debido a los límites del presente artículo, pero que demuestran cómo los problemas más generales de las humanidades repercuten también en las humanidades digitales.

Una de las cuestiones es la manera en la que la pedagogía digital —como se ha visto en las actuales polémicas sobre los cursos masivos abiertos en línea (MOOC), o sobre los «campus» totalmente digitales en las universidades públicas, etcétera— presenta la posibilidad de llevar a cabo cambios gigantescos en los objetivos, las prácticas, los destinatarios, la infraestructura y el personal de la enseñanza de las humanidades. El engaño está sustentado en décadas de legitimación defensiva de las humanidades como «habilidades» del pensamiento y del lenguaje añadidas a los conocimientos humanísticos tradicionalmente significativos. Si las humanidades encajan en parte en ese molde, entonces tal vez —según piensan algunos administradores y legisladores— deberían meterse totalmente en él para que su contenido pueda «entregarse» modularmente a través de Internet, a la manera de los MOOC o los cursos de la Academia Khan en ciencia, tecnología, ingeniería y matemáticas que han gozado de tanta popularidad entre el público.

Otra cuestión es el mercado laboral que se abre para los humanistas digitales, especialmente en lo que se refiere a las posibilidades de titularidad para puestos de profesor universitario, frente a la precariedad de los puestos de adjunto o asociado, de personal técnico, de personal de apoyo o puestos en instituciones culturales como opción alternativa a la carrera académica. El constante cambio y la incertidumbre que rodea a los trabajos relacionados con las humanidades digitales hacen que estos registren en realidad la tasa más alta de incertidumbre laboral en de las humanidades, puesto que los trabajos «importantes» quedan cada vez más lejos de ser los soñados puestos de titular en universidades y se tiende al ideal corporizado de convertirse en trabajadores con un conocimiento profesional o de gestión reconfigurable y reemplazable, debido a lo que están constantemente amenazados por la sombra de los procesos de reestructuración en favor de la externalización del trabajo y de trabajadores permanentemente temporales [«*permatemp*»] que están todavía más explotados.

Así pues, las humanidades digitales acusan de todas estas maneras la crisis de las humanidades. Por esa razón, algunos otros colegas y yo echamos a andar la iniciativa *4Humanities*, «impulsada por la comunidad de las humanidades digitales», para que las humanidades digitales puedan tratar de defender las humanidades y no solo acusar su crisis. Yo no sé hasta qué punto esta iniciativa y otras similares causarán o no impacto en la importancia que se les dé a las humanidades en el mundo. Pero sí sé que este esfuerzo de poner las humanidades digitales al servicio del alma de las humanidades es lo que ahora mismo tiene sentido para un humanista, digital o no digital<sup>20</sup>.

---

## NOTAS

20 | Defender las humanidades no significa necesariamente sólo defender las nociones más antiguas o disciplinarias de las humanidades. Mi trabajo «The Humanities and Tomorrow's Discoveries» (2012a) intenta reformular las humanidades haciendo causa común con otras disciplinas y con las necesidades públicas.

## Bibliografía citada

- ADAMS, M. (ed.) (2010): «*Cunning Passages, Contrived Corridors*»: *Unexpected Essays in the History of Lexicography*, Monza: Polimetrica.
- ALEXANDER, M. (2012): Mensaje al autor. 6 Oct. 2012. E-mail.
- ALEXANDER, M. (2010): «The Various Forms of Civilization Arranged in Chronological Strata»: Manipulating the HTOED» en Adams, M. (ed.), «*Cunning Passages, Contrived Corridors*»: *Unexpected Essays in the History of Lexicography*, Monza: Polimetrica, 309-323.
- «Analytics: Supervised vs. Unsupervised Learning». *MONK: Metadata Offer New Knowledge*, [01/12/2012].
- BAUER, J. (2011): «Who You Calling Untheoretical?», *Journal of Digital Humanities*, 1, 1, 68-74, [23/11/2012].
- BEARMAN, P. S., y STOVEL, K. (2000): «Becoming a Nazi: A Model for Narrative Networks», *Poetics*, 27, 69-90.
- «Bibliography» en Düring, M. (ed.), *Historical Network Research*. 18 Sept. 2012, [26/11/2012].
- BOGEN, J. y WOODWARD, J. (1988): «Saving the Phenomena», *Philosophical Review*, 97, 3, 303-352.
- BURDICK, A.; DRUCKER, J.; LUNENFELD, P.; PRESNER, T. y SCHNAPP, J. (2012): *Digital\_Humanities*, Cambridge: MIT Press.
- CECIRE, N. (2011a): «Introduction: Theory and the Virtues of Digital Humanities», *Journal of Digital Humanities*, 1, 1, 44-53, [23/11/2012].
- CECIRE, N. (2011b): «When Digital Humanities Was in Vogue», *Journal of Digital Humanities*, 1, 1, 54-58, [23/11/2012].
- «The Digital Humanities Manifiesto 2.0», *Humanities Blast*, 2009, [09/09/2012].
- FEYERABEND, P. (1978): *Against Method: Outline of an Anarchistic Theory of Knowledge*, Londres: Verso.
- FRANZOSI, R. (2010): *Quantitative Narrative Analysis*, Los Ángeles: Sage.
- GIBBS, F. W. y COHEN, D. J. (2011): «A Conversation with Data: Prospecting Victorian Words and Ideas», *Victorian Studies*, 54, 1, 69-77.
- GOLD, M. K. (ed.) (2012): *Debates in the Digital Humanities*, Minneapolis: University of Minnesota Press.
- GOLDBERG, D. T. y SVENSSON, P. (ed.) (2015): *Between Humanities and the Digital*, Cambridge: MIT Press.
- GOLDSTONE, A. y UNDERWOOD, T. (2012): «What Can Topic Models of PMLA Teach Us about the History of Literary Scholarship?», *The Stone and the Shell: Historical Questions Raised by a Quantitative Approach to Language*, 14 Dic. 2012, [09/12/2012].
- HAYLES, N. K. (2012): *How We Think: Digital Media and Contemporary Technogenesis*, Chicago: University of Chicago Press.
- HENSELER, C.; LIU, A.; ROCKWELL, G.; SINCLAIR, S. y TERRAS, M.; con THOMAS, L. (eds.). *4Humanities: Advocating for the Humanities*, 4Humanities, [16/01/2013].
- HEUSER, R. Mensaje al autor. 13 Sept. 2012. E-mail.
- HEUSER, R. y LE-KHAC, L. (2011): «Learning to Read Data: Bringing Out the Humanistic in the Digital Humanities», *Victorian Novels*, 54, 1, 79-86.
- HEUSER, R. y LE-KHAC, L. (2012a): «Online Companion to “Learning to Read Data: Bringing Out the Humanistic in the Digital Humanities”, *Victorian Studies*, 54, 1; y Stanford Literary Lab Pamphlet 4», *Stanford Literary Lab*. Stanford Lit. Lab, Mayo 2012, [13/09/2012].
- HEUSER, R. y LE-KHAC, L. (2012b): *A Quantitative Literary History of 2,958 Nineteenth-Century British Novels: The Semantic Cohort Method*. *Stanford Literary Lab*. Stanford Lit. Lab, Mayo 2012, [01/09/2012]. Pamphlet 4.
- Historical Thesaurus. Oxford English Dictionary*, Oxford: Oxford University Press, 2012, [25/11/2012].
- Historical Thesaurus Photo Gallery*, University of Glasgow, [25/09/2012].
- JOCKERS, M. L. (2013): «Thoughts on a Literary Lab», *Matthew L. Jockers*, 4 Enero 2013, [10/01/2013].
- JOCKERS, M. y WORTHEY, G. (2011): «Introduction: Welcome to the Big Tent», *Digital Humanities 2011*, [23/11/2012].

- KAY, C. (2010): «Classification: Principles and Practice» en Adams, M. (ed.), «*Cunning Passages, Contrived Corridors*»: *Unexpected Essays in the History of Lexicography*. Edición electrónica. Monza: Polimetrica, 255-270.
- KAY, C. (2012): «What Is the *Historical Thesaurus* of the OED», *Oxford English Dictionary*, Oxford: Oxford University Press, [25/11/2012].
- KAY, C.; ROBERTS, J.; SAMUELS, M. y WOTHERSPOON, I. (2009): *Historical Thesaurus of the Oxford English Dictionary, with Additional Material from «A Thesaurus of Old English»*, Oxford: Oxford University Press.
- KAY, C. y CHASE, T. J. P. (1987): «Constructing a Thesaurus Database», *Literary and Linguistic Computing*, 2, 3, 161-163.
- KIRSCHENBAUM, M. G. (2012): «Digital Humanities As/ Is a Tactical Term» en Gold, M. K. (ed.), *Debates in the Digital Humanities*, Minneapolis: University of Minnesota Press, 415-428.
- KIRSCHENBAUM, M. G. (2010): «What Is Digital Humanities and What's It Doing in English Departments?», *ADE Bulletin*, 150, 55-61.
- KIRSCHENBAUM, M. G. (2008): *Mechanisms: New Media and the Forensic Imagination*, Cambridge: MIT Press.
- KOH, A. (2012): «More Hack, Less Yack? Modularity, Theory and Habitus in the Digital Humanities», *Adeline Koh*, 21 Mayo 2012, [23/11/2012].
- LATOUR, B. (2005): *Reassembling the Social: An Introduction to Actor-Network-Theory*, Oxford: Oxford University Press.
- LE-KHAC, L. Mensaje al autor. 4 Feb. 2013. E-mail.
- LEMERCIER, C. y ROSENTAL, P.-A. (2010): «The Structure and Dynamics of Migration Patterns in Nineteenth-Century Northern France», Vers. 3, *HAL: Sciences de l'homme et de la société*. Centre pour la Communication Scientifique Directe, 5 Mayo 2010, [24/11/2012].
- LIU, A. (2013a): «From Reading to Social Computing» en Price, K. M. y Siemens, R. (eds.), *Literary Studies in the Digital Age: An Evolving Anthology. MLA Commons*. MLA.
- LIU, A. (2013b): «Is Digital Humanities a Field? An Answer from the Point of View of Language», *Alan Liu*, 6 Marzo 2013, [06/03/2013].
- LIU, A. (2012a): «The Humanities and Tomorrow's Discoveries», *4Humanities: Advocating for the Humanities*, 4Humanities, 14 Marzo 2013, [14/03/2013].
- LIU, A. (2012b): «The State of the Digital Humanities: A Report and a Critique», *Arts and Humanities in Higher Education*, 11, 1-2, 8-41.
- LIU, A. (2012c): «Where Is Cultural Criticism in the Digital Humanities?» en Gold, M. K. (ed.), *Debates in the Digital Humanities*, Minneapolis: University of Minnesota Press, 490-509.
- LIU, A. (2004): *The Laws of Cool: Knowledge Work and the Culture of Information*, Chicago: University of Chicago Press.
- MCCARTY, W. (2005): *Humanities Computing*, Basingstoke: Palgrave.
- MCGANN, J. J. (1999): «Position Statement from McGann» en Hockey, S., Renear, A. y McGann, J. J., *Panel: What Is Text? A Debate on the Philosophical and Epistemological Nature of Text in the Light of Humanities Computing Research*, [25/11/2012].
- MCGANN, J. J. (2001): *Radiant Textuality: Literature after the World Wide Web*, Nueva York: Palgrave.
- MOHR, J. W. y DUQUENNE, V. (1997): «The Duality of Culture and Practice: Poverty Relief in New York City, 1888–1917», *Theory and Society*, 26, 2-3, 305-356.
- MORETTI, F. (2011): *Network Theory, Plot Analysis. Stanford Literary Lab*, Stanford Lit. Lab, 1 Mayo 2011, [06/07/2012]. Pamphlet 2.
- «Names, Dates, People, and Places». *P5: Guidelines for Electronic Text Encoding and Interchange*, Vers. 2.2.0, TEI, 25 Oct. 2012, [14/01/2013].
- National Endowment for the Humanities, Office of Digital Humanities. *Digital Humanities Start-Up Grants: Frequently Asked Questions*, NEH, [17/11/ 2012].
- OWENS, J. B. «Jack» (2007): «What Historians Want from GIS», *ArcNews*, Esri, [15/01/2013].
- PHILLIPS, A. «#transformDH—a Call to Action Following ASA2011», *HASTAC*, 26 Oct. 2011, [24/11/2012].
- PICKERING, A. (1995): *The Mangle of Practice: Time, Agency, and Science*, Chicago: University of Chicago Press.

- PRICE, K. M. y SIEMENS, R. (eds.) (2013): *Literary Studies in the Digital Age: An Evolving Anthology*. *MLA Commons*. MLA, [12/01/2013].
- RAMSAY, S. (2008): «Algorithmic Criticism» en Schreibman, S. y Siemens, R. (eds.), *A Companion to Digital Literary Studies*, Oxford: Blackwell, *Alliance of Digital Humanities Organizations*, [25/11/2012].
- RAMSAY, S. (2011a): «On Building», *Stephen Ramsay*, 11 Ene. 2011, [20/06/2011].
- RAMSAY, S. (2011b): «Who's In and Who's Out», *Stephen Ramsay*, 8 Ene. 2011, [23/11/2012].
- SCHREIBMAN, S.; SIEMENS, R. y UNSWORTH, J. (eds.) (2004): *A Companion to Digital Humanities*, Oxford: Blackwell, *Alliance of Digital Humanities Organizations*, [27/11/2012].
- SCULLEY, D. y PASANEK, B. M. (2008): «Meaning and Mining: The Impact of Implicit Assumptions in Data Mining for the Humanities», *Literary and Linguistic Computing*, 23, 4, 409-424.
- SEEFELDT, D. y THOMAS III, W. G. (2009): «What Is Digital History? A Look at Some Exemplar Projects», *Faculty Publications*, Dept. of History, U of Nebraska, Lincoln, 1 Mayo 2009, [12/10/ 2012]. Paper 98.
- STALLYBRASS, P. (2002): «Books and Scrolls: Navigating the Bible» en Andersen, J. y Sauer, E. (eds.), *Books and Readers in Early Modern England: Material Studies*, Filadelfia: University of Pennsylvania Press, 42-79.
- «The Structure of the Historical Thesaurus of the OED», Oxford University Press, 2009, [25/11/2012].
- SVENSSON, P. (2012): «Envisioning the Digital Humanities», *Digital Humanities Quarterly*, 6, 1, [23/11/2012].
- SVENSSON, P. (2010): «The Landscape of Digital Humanities», *Digital Humanities Quarterly*, 4, 1, [23/11/2012].
- SVENSSON, P. (2009): «Humanities Computing as Digital Humanities», *Digital Humanities Quarterly*, 3, 3, [30/11/2010].
- UNDERWOOD, T. (2012): «Topic Modeling Made Just Simple Enough», *The Stone and the Shell: Historical Questions Raised by a Quantitative Approach to Language*, 7 Abr. 2012, [14/01/2013].
- UNDERWOOD, T. y SELLERS, J. (2012): «The Emergence of Literary Diction», *Journal of Digital Humanities*, 1, 2, [04/02/2013].
- VISMANN, C. (2008): *Files: Law and Media Technology*, Winthrop-Young, G. (trad.), Stanford: Stanford University Press.
- WOTHERSPOON, I. (1992): «Historical Thesaurus Database Using Ingres», *Literary and Linguistic Computing*, 7, 4, 218-225.
- WOTHERSPOON, I. (2010): «The Making of The Historical Thesaurus of the Oxford English Dictionary» en Adams, M. (ed.), «*Cunning Passages, Contrived Corridors*»: *Unexpected Essays in the History of Lexicography*, Monza: Polimetrica, 271-287.