

CODIFICACIÓN Y REPRESENTACIÓN CARTOGRÁFICA DE NOTICIAS. APLICACIÓN DE LAS HUMANIDADES DIGITALES AL ESTUDIO DEL PERIODISMO DE LA EDAD MODERNA

Francisco Baena-Sánchez, Carlota Fernández-Travieso, Carmen Espejo-Cala y Javier Díaz-Noci



Francisco Baena-Sánchez es profesor del *Departamento de Periodismo 1* de la *Universidad de Sevilla*. Especializado en historia del periodismo, ha publicado estudios sobre la prensa de la Edad moderna, sobre la propaganda obrera en Andalucía y sobre la prensa local de Huelva, Cádiz y Sevilla. Dirige la revista de información universitaria *NexUS* y es miembro del proyecto de I+D *Biblioteca digital Siglo de Oro*. Ha sido vicedecano de estudiantes de la *Facultad de Comunicación* y subdirector de la revista *I/C. Revista universitaria en información y comunicación*.
<http://orcid.org/0000-0002-5304-7924>

Universidad de Sevilla, Departamento de Periodismo 1, Facultad de Comunicación
Av. Américo Vespucio, s/n. 41092 Sevilla, España
frbaena@us.es



Carlota Fernández-Travieso, investigadora post-doctoral en la *Universidade da Coruña* y miembro del *Sielae* –grupo pionero en la aplicación de las humanidades digitales en España–, colabora desde 2004 en la alimentación y mejoras técnicas del *Catálogo y biblioteca digital de relaciones de sucesos*. Con el objetivo de ampliar esta línea de investigación, adquirió conocimientos sobre la *TEI* en la *Brown University* y, tras diseñar un esquema que permita codificar relaciones y gacetas, planea, junto a otros miembros de las universidades de *A Coruña* y *Sevilla*, iniciar la marcación sistemática de los textos.

<http://orcid.org/0000-0002-0882-2688>

Universidade da Coruña, Departamento de Filología Española e Latina, Facultade de Filoloxía
Campus da Zapateira, s/n. 15071 A Coruña, España
cfernandez@udc.es



Carmen Espejo-Cala es profesora titular del *Departamento de Periodismo I, Universidad de Sevilla*, especializada en historia del periodismo de la Edad moderna. Directora del *Grupo Paidi Historia del periodismo y las lecturas populares en Andalucía*, e investigadora en el proyecto de I+D *Bibliotecas digitales Siglo de Oro*. Ha sido investigadora en la red europea *News Networks in Early Modern Europe* (2011-13). Ha sido coordinadora del *Máster universitario en comunicación y cultura* y directora de la revista *I/C. Revista universitaria en información y comunicación*. Coordinó el curso *Barroco y comunicación* (2009), dirigido por Roger Chartier.

<http://orcid.org/0000-0001-9949-5901>

Universidad de Sevilla, Departamento de Periodismo 1, Facultad de Comunicación
Av. Américo Vespucio, s/n. 41092 Sevilla, España
carmenes@us.es



Javier Díaz-Noci, doctor en historia, es catedrático del *Departamento de Comunicación* de la *Universitat Pompeu Fabra (UPF)* de Barcelona. Ha sido profesor visitante de las *Universidades de Oxford, Nevada* en Reno (EUA) y *Federal de Bahía* (Brasil), entre otras. Coordinador del *Máster universitario en comunicación social* de la *UPF*. Es miembro del grupo de investigación *Audiencias activas y periodismo. Interactividad, integración en la Web y buscabilidad de la información periodística*. CSO2012-39518-C04-02. *Plan Nacional de I+D+i, Ministerio de Economía y Competitividad* (España).

<http://orcid.org/0000-0001-9559-4283>

Universitat Pompeu Fabra, Departamento de Comunicación, Facultad de Comunicación
Roc Boronat, 138. 08018 Barcelona, España
javier.diaz@upf.edu

Resumen

Las humanidades digitales ofrecen nuevas posibilidades para el estudio de la historia del periodismo de la Edad moderna. Desde la perspectiva de la *digital history*, planteamos la necesidad de buscar estrategias unificadas que empleen los estándares y protocolos más extendidos en la comunidad internacional, para así poder construir *corpora* marcados digitalmente y susceptibles de ser analizados de forma sistemática. En el artículo se revisan los criterios que tradicionalmente se han utilizado para describir las gacetas a partir del concepto de serialidad. A continuación, se presenta una propuesta personalizada de esquema xml/TEI para codificar este tipo de textos periodísticos. También se exponen los resultados de un estudio parcial sobre la representación cartográfica en internet de las redes de información que existían en la Edad moderna. Finalmente, las conclusiones revelan la pertinencia y la utilidad de aplicar las diferentes herramientas de las humanidades digitales al estudio del primer periodismo europeo.

Palabras clave

Humanidades digitales, Historia digital, Historia del periodismo, Edad moderna, Gaceta, Relación de sucesos, Periódicos, Marcado xml/TEI, Representación cartográfica de datos.

Title: Encoding and mapping news. Applying digital humanities methods to research on journalism in the Modern age

Abstract

Digital humanities offer new possibilities to do historical research on journalism in the Modern age. Based on the so-called digital history techniques and methods, we suggest a need to find new, uniform strategies that use the most commonly used standards and protocols in the international scholarly community, so we can construct digitally tagged *corpora* for future systematic analysis. We review the traditional criteria used to describe newssheets, using the concept of seriality and propose a customized xml/TEI schema to codify this type of texts. We present the preliminary results of the mapping of the news flow networks in the Modern age in Europe. Finally, the conclusions shed some light on the relevance and benefits of applying digital humanities tools to the study of the first European manifestations of what was to become known as 'journalism'.

Keywords

Digital humanities, Digital history, Journalism history, Modern age, Newsbook, Newssheets, Newspapers, Periodicals, xml/TEI, Markup, Mapping, Cartographic.

Baena-Sánchez, Francisco; Fernández-Travieso, Carlota; Espejo-Cala, Carmen; Díaz-Noci, Javier (2014). "Codificación y representación cartográfica de noticias. Aplicación de las humanidades digitales al estudio del periodismo de la Edad moderna". *El profesional de la información*, septiembre-octubre, v. 23, n. 5, pp. 519-526.

<http://dx.doi.org/10.3145/epi.2014.sep.09>

1. Introducción

Durante la Edad moderna España y Portugal ocuparon una posición geográfica periférica en relación con los otros países europeos, pero, a la vez, central en la recepción, distribución y producción de noticias. Sin embargo, la historia del primer periodismo español debe ser profundamente revisada, y es mucho más rica y compleja de lo que la importante, pero necesariamente limitada, historiografía positivista había concluido (Hartzenbusch, 1873; Varela-Hervías, 1960). Para comenzar, existieron publicaciones periódicas antes de 1661, año de publicación de la *Gazeta nueva* de Madrid, como han demostrado los trabajos de Carmen Espejo (2006, 2013 a y b; Guillamet, 2003), y otros. En segundo lugar, la historia del periodismo no debe reducirse al estudio de la prensa periódica. En tercer lugar, la prensa en lengua castellana se produjo también fuera de la Península. Sobre todo: necesitamos cuestionar cuál era la posición y los mecanismos de los flujos noticiosos (Espejo, 2013a) en lengua castellana (y en otras lenguas peninsulares, singularmente el portugués y el catalán), cuáles los focos informativos y cuáles las estrategias de

producción empleadas por quienes estaban desarrollando el primer periodismo de la época moderna.

Para ello, necesitamos todas las herramientas posibles de las denominadas humanidades digitales al alcance y al servicio de la investigación histórica, y en concreto del área a que nos referimos, que se ha denominado *digital history* (Cohen; Rosenzweig, 2005) así como contemplar aspectos de preservación digital de documentos antiguos (Higgs, 1998; Deegan; Tanner, 2006). Tanto para la catalogación del ingente material, en buena parte aún desconocido o poco sistematizado y explotado, que puede desembocar no sólo en catálogos de tipo censal (Salmon Crane; Kaye, 1927; Nelson; Secombe, 1987; Domínguez-Guzmán, 1992), en diccionarios de publicaciones cuyo modelo serían los coordinados para el caso francés por Jean Sgard (1991; 1992); (Espejo, 2008), como para la construcción de *corpora* marcados digitalmente y susceptibles de ser analizados de forma sistemática en forma de *corpus-drive analysis* (y no meramente *corpus-based analysis*), tal como puso de manifiesto Elena Tognini-Bonelli (2001).

Tenemos que señalar también la necesidad de actuar en interlocución con colegas de otros países, en busca de una perspectiva paneuropea, siguiendo la terminología de **Henry Ettinghausen** (2013). Resulta indispensable mantener un diálogo con investigadores de toda Europa para establecer una red estable (ya iniciada por el proyecto *News networks in Early Modern Europe*, liderado por el británico **Joad Raymond**), y favorecer el intercambio de conocimientos pero también, y en ese aspecto nos centramos en este texto, de protocolos técnicos estándar, tanto para establecer la tipología de impresos y manuscritos a estudiar, como para catalogarlos de forma unitaria y, sobre todo, acceder de forma semejante y en condiciones óptimas de explotación al contenido de esos productos informativos. En España disponemos del ingente esfuerzo llevado a cabo por el grupo encabezado por las profesoras de la *Universidad de La Coruña* Sagrario López-Poza y Nieves Pena-Sueiro, sobre relaciones de noticias, que técnicamente también es un modelo a seguir.

Por lo que respecta a la constitución de *corpora*, necesitamos no sólo una mera imagen facsimilar de los documentos, importante sin ningún género de dudas, sino también transcripciones en forma de documentos buscables, “un gran y estructurado conjunto de textos que puedan ser almacenados y analizados electrónicamente” (**Tognini-Bonelli**, 2001, p. 65), un *machine-readable set of texts*. Si quisiéramos llevar a cabo algún tipo de análisis cuantitativo del discurso, disponemos de dos magníficos ejemplos en países vecinos, el de **Stéphane Haffemayer** (2002) en Francia y el del grupo liderado por **Pedro-Jorge Sousa** (2011; 2012) en Portugal puesto que ambos han trabajado sobre la base de frecuencias y concordancias de palabras y frases. Existen también otros modelos con los que relacionarse, tales como el *ZEN* (*Zurich English newspaper corpus*) de Fries y Schneider, que cubre el período que va desde 1661 hasta 1791, y el *FEEN* (*Florence early English newspapers*) en el *Lancaster news-book corpus*.

<http://www.es.uzh.ch/Subsites/Projects/zencorpus.html>
<http://cqpweb.lancs.ac.uk>

“Necesitamos las herramientas de las humanidades digitales al servicio de la historia del periodismo, tanto para la catalogación como para la construcción de *corpora* marcados digitalmente y susceptibles de ser analizados de forma sistemática”

No podemos, por otra parte, soslayar el desigual desarrollo de los proyectos, como por ejemplo el de los fondos hemerográficos conservados en la *Biblioteca Nacional de España*, en la *Hemeroteca Municipal de Madrid* o en instancia europeas que deberían reunir y unificar los esfuerzos nacionales, como la biblioteca digital *Europeana*. Pero es aún más grave que, a los problemas historiográficos derivados de la existencia de colecciones fragmentarias, deficientemente

catalogadas, incompletas, se una la ausencia de estrategias unificadas que empleen los estándares y protocolos más extendidos en la comunidad internacional: *OAI-PMH* (*Open archive initiative-protocol for metadata harvesting*), *ALTO* (*analyzed layout and text object*) –una extensión de *METS* (*metadata encoding and transmission schema*) que mantiene el aspecto facsimilar del documento y a la vez indica y coordina los elementos de la página-, y *SKOS* (*simple knowledge organization system*).

La constitución de un corpus de estas características no es un trabajo fácil. Partiendo de un facsímil en formato de imagen o pdf, la transcripción del texto original puede ser difícilmente confiada en exclusiva a un programa de *OCR* (*optical character recognition*), que no siempre funciona correctamente con este tipo de documentos, por lo que necesariamente debe intervenir manualmente. Necesitamos desarrollar estrategias flexibles de conversión de textos, pero correctamente diseñadas desde el principio. Esto incluye marcar digitalmente el texto incorporando *tags* del lenguaje xml (*extended markup language*) y produciendo metadatos, por no mencionar la necesidad de traducir todos los textos a una única lengua si queremos compararlos con otros *corpora* y aplicarles diferente software de análisis, como *WMatrix* o *CopyCatch*.

“A los problemas historiográficos por las colecciones deficientemente catalogadas o incompletas, se une la ausencia de estrategias unificadas que empleen los estándares y protocolos más extendidos en la comunidad internacional”

2. Gacetas seriadas: una revisión metodológica de los criterios de descripción

Como se ha señalado ya, el nacimiento del periodismo en Europa está mal contado; particularmente, la gaceta –impreso que recopilaba noticias de la actualidad internacional, previo a la aparición del diario– resulta ser un género deficitariamente conocido y mal comprendido, y por ello a menudo queda fuera de las historias del periodismo o relegado a un breve capítulo inicial plagado de generalidades y errores. Los criterios contemporáneos para la definición del periódico se aplican a estos productos impresos que fueron creados en otras condiciones de mercado, y que por tanto casi nunca presentan los rasgos que permitirían una fácil adscripción a la tipología de los periódicos. Enumeramos algunos de estos criterios conflictivos:

– Periodicidad. Aparece muy tardíamente en el mercado periodístico. Había noticias impresas y público demandante de las mismas antes de las gacetas periódicas, que no se consolidan en Europa hasta mediados del XVII. Pero, además, esa consolidación de la periodicidad no se produce de manera idéntica en toda Europa ni tampoco es perfecta o exacta salvo excepciones. De hecho, un fenómeno al que se ha prestado hasta ahora una atención

insuficiente es el de los productos periodísticos seriados (pero todavía no periódicos), que vivieron una verdadera explosión en el mercado español en las cuatro o cinco primeras décadas del XVII. Pero esta “serialidad”, previa a la periodización, es difícil de detectar en la descripción de un impreso, de manera que las bases de datos o bibliotecas consignan cada uno de los números como una publicación independiente. Esto justifica que siga afirmándose en manuales y monografías que la prensa periódica arranca en España más tarde que en el resto de Europa.

- Cabecera. Una dificultad derivada de la anterior es que estas series carecen de una cabecera estable en el tiempo. Aunque una denominación que alude al género del impreso suele estar siempre presente, la denominación no es siempre la misma. Así, en la misma serie, podemos encontrar números que se titulan “gaceta” y otros que se titulan “relación”, “carta” o “avisos”. En este caso también, las bases de datos y bibliotecas registran cada número como publicación independiente.
- Diseño y estructura. Todas estas deficiencias en la descripción derivan además de una carencia historiográfica: los historiadores del periodismo desconocemos aún muchos rasgos del mercado periodístico de la Edad moderna, de manera que no sabemos cuál era el público real de estos impresos y, sobre todo, qué rasgos hacían reconocible y apetecible para el público a estos primeros periódicos.

Nuestra hipótesis es que el diseño y la estructura, además de algunos rasgos constantes en el contenido de la portada, funcionaban ya como elementos que hacían reconocible para el lector la serialidad de la gaceta, a pesar de las variaciones en la cabecera o la falta de numeración. Un análisis de contenido que utilice las etiquetas oportunas podría llevarnos a reconocer estas regularidades que hoy pasan desapercibidas al investigador, gracias a la capacidad de las humanidades digitales para trabajar con series extensas de documentos.

Este trabajo tiene como objetivo principal comprobar la pertinencia y utilidad de aplicar recursos de las humanidades digitales al estudio del periodismo de la Edad moderna. Específicamente, se plantea la necesidad de personalizar un esquema de etiquetas que permita marcar fehacientemente las características de estos textos periodísticos antiguos. Asimismo, se pretende probar la aplicación de las nuevas plataformas de creación y publicación de información geográfica en internet al estudio de la historia del periodismo moderno.

3. Propuesta de un esquema xml-TEI para la codificación de relaciones y gacetas

A la hora de crear un corpus de relaciones y gacetas que puedan ser almacenadas y analizadas electrónicamente, resulta ineludible valorar la aportación de las directrices de la TEI (*Text encoding initiative*), que tantos proyectos –como el ZEN corpus y los trabajos de Brownlees previamente mencionados o los aludidos en la web de la propia TEI– han utilizado hasta la fecha.

<http://www.tei-c.org/Activities/Projects>

Basándose en el lenguaje de marcas xml, la TEI nos ofrece un esquema que incluye y describe las etiquetas necesarias

para marcar las características estructurales, interpretativas y conceptuales de los tipos de textos más frecuentes en humanidades. Este esquema provee etiquetas para marcar los metadatos y los distintos elementos que la mayoría de los textos contiene (párrafos, versos, columnas, páginas, imágenes, encabezados, etc.), pero también describe módulos especializados, por ejemplo, para codificar obras teatrales o describir manuscritos. De entre todos estos elementos, con la salvedad de un pequeño grupo que son obligatorios, las directrices de la TEI permiten seleccionar los que nos convengan, modificar sus características o incluso añadir otros para crear nuevos esquemas que sirvan a nuestros propósitos.

La capacidad de las humanidades digitales para trabajar con series extensas de documentos podría llevarnos a reconocer estas regularidades que hoy pasan desapercibidas al investigador

Así pues, la TEI no sólo nos proporciona un vocabulario en común con muchos otros investigadores, que facilita el intercambio de información, sino que gracias a esta posibilidad de *customizar* o personalizar su esquema nos permite llevar a cabo la edición digital transcrita de los textos de la prensa de la Edad moderna representando fehacientemente su estructura, incorporando un análisis que considera las diferentes categorías conceptuales por las que se interesa la historia del periodismo e indicando la relación entre los textos que hacen parte de una serie.

Al elaborar una propuesta de esquema –para la que se ha tomado como base el ingente corpus de textos que reproduce facsimilarmente el *Catálogo y biblioteca digital de relaciones de sucesos*–, hemos hallado dificultades para usar las etiquetas propias de la portada, pues, por su carácter restrictivo, no posibilitan la marcación de la información que, con frecuencia, se condensa en la primera página de nuestros documentos. Así, por ejemplo, la mención al autor o el año de producción pueden ser incluidas en el título, pero el elemento <titlePart>, que en TEI se utiliza para marcar el título, no permite la inclusión de <docAuthor> y <docDate>. En nuestros textos frecuentemente aparecen otras informaciones (como, por ejemplo, la dedicatoria) entremezcladas con otros datos que según la TEI son propios del elemento <titlePage>, sin embargo este elemento no permite la expresión <div type=“dedication”> en su interior. Así mismo, resulta común que la mención de que el texto cuenta con licencia se exprese junto a otros datos de impresión y no de modo independiente, pero la TEI no prevé que <docImprint> pueda incluir la etiqueta <imprimatur>. Estos conflictos, de los que se proporcionan ejemplos en el trabajo de Carlota Fernández-Travieso (2013), nos lleva a proponer, ahora, tras muchas hipótesis de trabajo previas, la utilización de etiquetas como <div type=“title”>, <div type=“docImprint”> o <div type=“imprimatur”> directamente dentro del elemento <front>, que enmarcaría todo contenido previo al inicio del texto propiamente dicho, o de las expresiones <name role=“author”> y <date type=“docDa-

te">, para marcar el nombre del autor y la fecha de edición. Esta solución nos permite, además, mejorar la descripción bibliográfica de los productos de la época de la imprenta manual, acuñando de modo similar la etiqueta <div type="privilege">, para marcar el privilegio, para el que la TEI no dispone de ninguna etiqueta.

Para marcar la información propia del género desde la perspectiva de nuestro interés, necesitamos crear vocabularios controlados para los valores de los atributos @type, @subtype y @role asociados a los elementos div (utilizado en TEI para marcar subdivisiones del texto), seg (para cualquier división del texto inferior a las divisiones autónomas) name (para marcar fechas), place (para marcar lugares) y date (para marcar fechas), elementos seleccionados por su gran flexibilidad para aparecer en distintas partes del texto. Así proponemos utilizar expresiones como, por ejemplo, <div type="news" n="01">, para marcar cada noticia, ya sea una o más las que contiene el texto, que a su vez podríamos contar gracias al atributo @n; <div type="translation"> para marcar la indicación de que el texto es una traducción de otro; <div type="continuation">, la indicación de que un texto tiene continuación; <seg type="docName">, el nombre que se da al texto: relación, gaceta, aviso, carta...; <seg type="infoSource">, su fuente; <place type="infoOrig">, el lugar de donde proviene una información; y tantas etiquetas más como precisemos para después recoger sistemática y fácilmente la información para apuntalar nuestros estudios. De manera similar, se utilizarán las etiquetas <name role="publisher"> para señalar sólo al costeador de la edición, y <name role="printer"> para señalar su impresor, distinguiendo así dos papeles que la TEI obvia al proveer sólo la etiqueta <publisher>.

La posibilidad de *customizar* el esquema TEI nos permite representar las particularidades estructurales de los textos periodísticos así como marcar su carácter serial

Los textos que pertenezcan a una serie incorporarán la etiqueta <SeriesStmnt> en los metadatos. El valor del atributo @xml:id aplicado a este elemento coincidirá para toda la serie, funcionando así como identificador de la misma; mientras que el atributo @n nos permitirá otorgar un orden

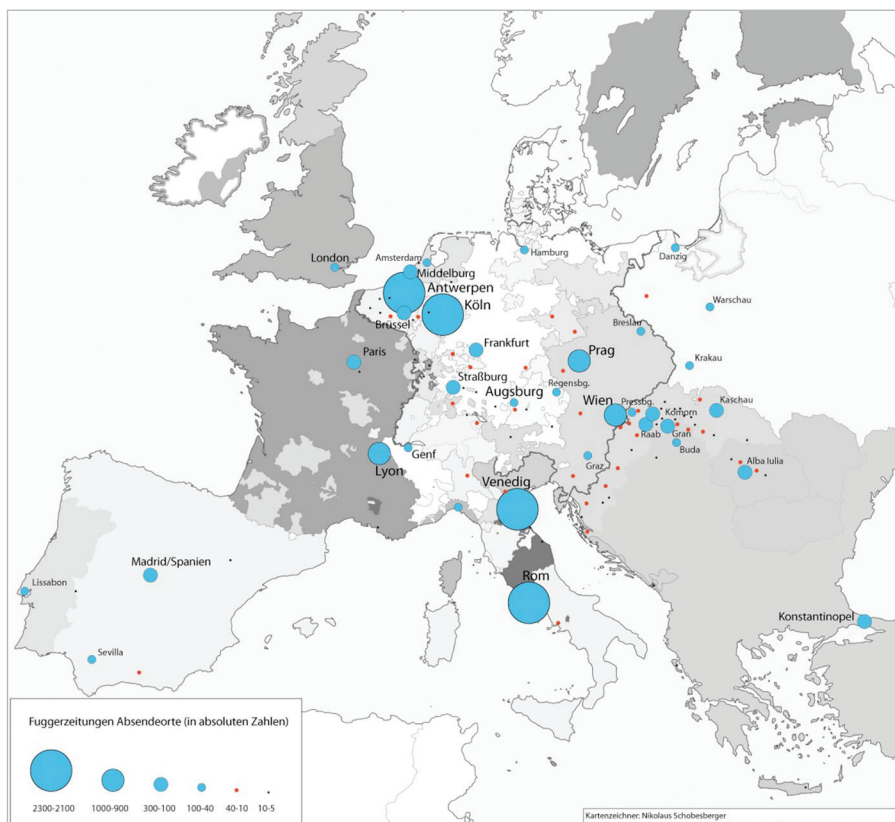


Figura 1. Lugares de procedencia de los periódicos incluidos en la gaceta vienesa *Fuggerzeitungen*. Fuente: **Nikolaus Schobesberger**, *The Fuggerzeitungen*. http://www.univie.ac.at/fuggerzeitungen/en/?page_id=336

dentro del conjunto a cada texto. Con esta solución lograríamos representar la relación de serialidad entre los documentos con suficiente profundidad como para diseñar una aplicación web que permita a los usuarios navegar fácilmente por la secuencia de textos etiquetados pertenecientes a una misma serie.

4. Una aproximación a la representación cartográfica digital del primer periodismo

En los últimos años, varios historiadores han utilizado la metáfora de la red para referirse al periodismo de la Edad moderna (**Raymond**, 2005; **Arblaster**, 2005; **Espejo**, 2012), una red a escala europea, cruzada por innumerables puntos de encuentro y de fuga, que precisa ser sobreimpresionada en un mapa (**Raymond**, 2012). Idealmente este mapa mostraría cómo circulaban las noticias en el complejo y dinámico mercado informativo de la Europa de los siglos XVI-XVII, esto es, tendría en cuenta los lugares donde se generaba y consumía la información, los canales por los que circulaba, la velocidad a la que viajaba y su desarrollo cronológico (un ejemplo aplicado al caso español en **Díaz-Noci**, 2012).

El trazado definitivo de ese mapa del primer periodismo europeo está cada vez más próximo gracias a las aportaciones que varios proyectos han realizado en los últimos años, siempre sobre la base de la convergencia entre la historia del periodismo y las humanidades digitales. Es el caso de *The Fuggerzeitungen. An early modern informative medium and its indexing*, en Austria y Alemania, y de *La gazette de Renaudot*, de **Stéphane Haffemayer**, en Francia. Una de sus

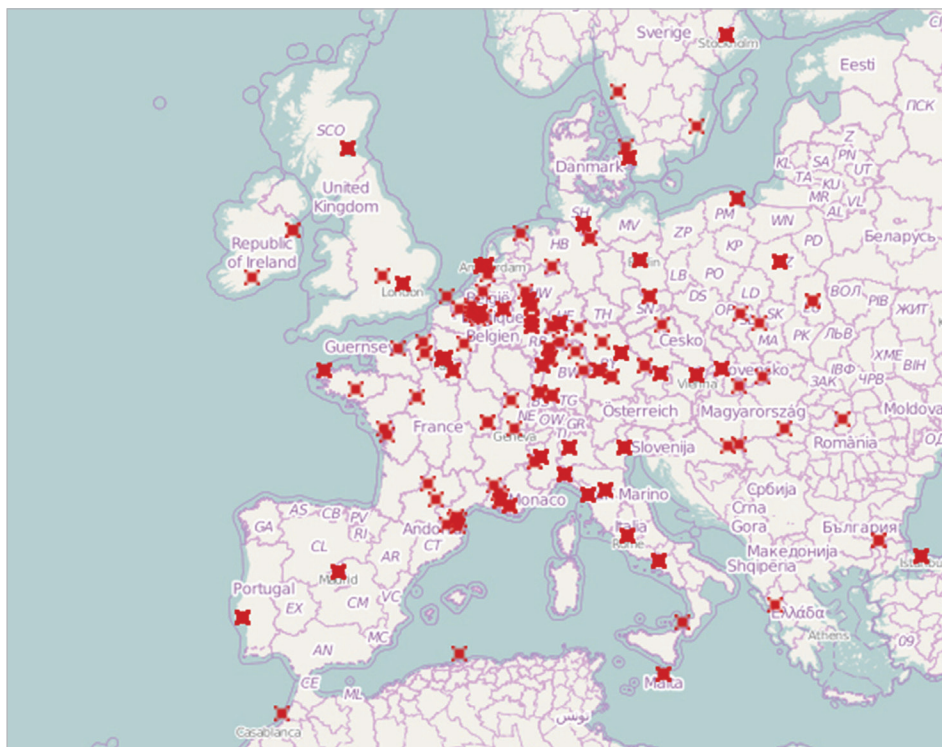


Figura 2. Lugares de procedencia de los avisos publicados por la *Gazette* en 1683, 1685 y 1689. Fuente: Stéphane Haffemayer, *La gazette de Renaudot*. <http://www.unicaen.fr/gazette/cartes.php>

aportaciones más originales consiste precisamente en el análisis cartográfico mediante mapas de la circulación de las noticias durante la Edad moderna (figuras 1 y 2).

Otro proyecto que ha combinado las posibilidades de la minería de textos con la representación cartográfica de datos para crear mapas interactivos es *Mapping the republic of letters*, en el mismo período histórico de la Edad moderna, pero fuera del ámbito de la historia del periodismo. El resultado ha sido una herramienta interactiva de análisis visual realizada con *Improvise*, un software libre distribuido bajo la *GNU General public license (GPL)*.

<http://republicofletters.stanford.edu/index.html#>

<http://www.stanford.edu/group/toolingup/rplviz/rplviz.swf>

El trazado definitivo del mapa del primer periodismo europeo está cada vez más próximo gracias a la convergencia que se ha producido entre la historia del periodismo y las humanidades digitales en los últimos años

En estos últimos años estamos asistiendo en el ámbito de los SIG (sistemas de información geográfica) a la aparición de una serie de plataformas que, a partir de datos de cualquier tipo asociados a una localización geográfica concreta, permiten crear, almacenar y compartir mapas interactivos en la nube de un modo asequible para usuarios inexpertos. Algunas de estas plataformas web, relativamente recientes y aún poco conocidas en el mundo académico, son *CartoDB* o *Geocommons*.

<http://cartodb.com>

<http://geocommons.com>

A continuación se muestran sucintamente los resultados de un estudio parcial que surge de la confluencia entre la necesidad de representar cartográficamente el periodismo de la Edad moderna y las posibilidades que ofrecen las plataformas web de publicación de información geográfica. En concreto, hemos utilizado la plataforma *CartoDB* con el fin de colaborar parcialmente a trazar ese mapa general acerca de la circulación de las noticias durante la Edad moderna.

Nuestra muestra está conformada por 36 gacetas publicadas en Sevilla por cinco impresores (Rodrigo de Cabrera, Juan Serrano de Vargas, Francisco de Lyra, Simón Fajardo y Juan de Cabrera) entre 1597 y 1629. Las ediciones digitaliza-

das, de acceso público, se conservan en el *Fondo antiguo de la Universidad de Sevilla* y en el *Catálogo y biblioteca digital Siglo de oro*. Específicamente, nuestro mapa (figura 3) representa la información geográfica contenida en la etiqueta <place type="infoOrig"> del esquema *TEI* descrito anteriormente, la cual nos permite marcar la procedencia de las fuentes de información (por ejemplo, avisos venidos desde Flandes), que los impresores sevillanos utilizaban para confeccionar sus gacetas.

<http://fondosdigitales.us.es/fondos>

<http://www.bidiso.es/RelacionesSucesosBusqueda>

5. Conclusiones

El ingente esfuerzo realizado en los últimos años, encaminado a la elaboración de bases de datos y hemerotecas digitales, ha favorecido el estudio de la prensa de la Edad moderna. Sin embargo, urge también dar el paso hacia una nueva etapa, dedicada al análisis exhaustivo de estos impresos periodísticos. Además, dada su condición de productos seriados, muchas veces no claramente periódicos, los investigadores debemos revisar los criterios de catalogación, descripción y análisis, para superar la deformación que supone aplicar criterios actuales a productos nacidos en mercados con características distintas.

Gracias a la posibilidad de *customizar* el esquema *TEI*, podremos crear un nuevo *relax ng* (un esquema de xml) que atienda a las particularidades estructurales de relaciones y gacetas, a las necesidades de analizar los textos desde la perspectiva de la historia del periodismo y al marcado de su carácter serial y que nos permita dar inicio a la creación de un corpus de textos que nos ayude a conocer mejor el conjunto de la prensa noticiosa de la Edad moderna y el origen del periodismo en España.

Fernández-Travieso, Carlota (2013). *Estudio de codificación xml/TEI para relaciones de sucesos españolas*. A Coruña: Sialae. ISBN: 9788469590416

<http://www.bidiso.es/sialae/upload/estaticas/file/FTCXMLEI1.pdf>

Guillamet, Jaume (2003). *Els orígens de la premsa a Catalunya. Catàleg de periòdics antics (1641-1833)*. Barcelona: Ajuntament de Barcelona. ISBN: 9788476096703

Haffemayer, Stéphane (2002). *L'information dans la France du XVII^e siècle. La gazette de Renaudot de 1647 à 1663*. Paris: Honoré Champion. ISBN: 2745304488.

Hartzenbusch, Eugenio (1873). *Apuntes para un catálogo de periódicos madrileños*. Madrid: Biblioteca Nacional de España.

Higgs, Edward (ed.) (1998). *History and electronic artefacts*. Oxford: Oxford University Press. ISBN: 9780198236337

Nelson, Carolyn; Seccombe, Matthew (1987). *British newspapers and periodicals 1641-1700*. New York: Modern Language Association of America. ISBN: 9780873521741

Raymond, Joad (2005). "Introduction: networks, communications, practice". *Media history*, v. 11, n. 1-2, pp. 3-19. <http://dx.doi.org/10.1080/1368880052000342389>

Raymond, Joad (2012). "El rostro europeo del periodismo inglés". En: Chartier, Roger; Espejo, Carmen (eds.). *La aparición del periodismo en Europa. Comunicación y propaganda en el Barroco*. Madrid: Marcial Pons, pp. 177-206. ISBN: 9788492820672

Salmon-Crane, Ronald; Kaye, Frederick B. (1927). *Census of British newspapers and periodicals 1620-1800*. Chapel Hill: University of North Carolina Press. ISBN: 9781469609416

Schobesberger, Nikolaus (2013). *Cartographical Analysis of the Viennese Fuggerzeitungen* http://www.univie.ac.at/fuggerzeitungen/en/?page_id=336

Sgard, Jean (ed.) (1991). *Dictionnaire des journaux (1600-1789)*. Paris; Oxford: Voltaire Foundation. ISBN: 9782740000045

Sgard, Jean (ed.) (1992). *Dictionnaire des journalistes (1600-1789)*. Paris; Oxford: Voltaire Foundation. ISBN: 9780729405386

Sousa, Pedro-Jorge et al. (2011). *A gazeta "da Restauração": Primeiro periódico português - Uma análise do discurso*, 2 vol. LabCom: Covilhã. ISBN: 9789896540609 <http://www.livroslabcom.ubi.pt/book/18>

Sousa, Pedro-Jorge et al. (2012). *Estudos sobre o Mercúrio português (1663-1667). Discurso e contexto*. LabCom: Covilhã. ISBN: 9789896540999 <http://www.livroslabcom.ubi.pt/book/93>

Tognini-Bonelli, Elena (2001). *Corpus linguistics at work*. Amsterdam; Philadelphia. ISBN: 9781588110619

Varela-Hervías, Eulogio (1960). *Gazeta nueva, 1661-1663. Notas sobre la historia del periodismo español en la segunda mitad del siglo XVIII*. Murcia: Nogués.

Próximos temas centrales

Noviembre 2014	Big data y analítica web
Enero 2015	Marketing de productos de información
Marzo 2015	Web semántica
Mayo 2015	Privacidad y seguridad de información
Julio 2015	Bases de datos bibliográficas
Septiembre 2015	Difusión científica en internet
Noviembre 2015	Periodismo computacional

Los interesados por favor consulten detenidamente las Normas para autores:

<http://www.elprofesionaldelainformacion.com/autores.html>

y luego envíen sus artículos a través del gestor de manuscritos OJS de la plataforma del Recyt:

<http://recyt.fecyt.es/index.php/EPI/index>