

Géneros web: líneas de investigación

Por Michela Montesi

Resumen: Se presenta una panorámica de las principales tendencias de la investigación sobre los géneros o tipos de documentos web. Los géneros son documentos determinados por los objetivos comunicativos que persiguen sus productores y receptores, así como por su forma y contenido. Conceptualizar los contenidos web en términos de géneros ofrece la posibilidad de seleccionarlos según un criterio complementario con respecto al tema o materia, susceptible además de implementación en sistemas automáticos de recuperación de información. Se propone una definición de géneros web, un resumen de los adelantos en su clasificación automática, y una perspectiva para organizar las cada vez más numerosas contribuciones sobre el tema.

Palabras clave: Géneros web, Tipos de documentos, Formas lingüísticas, Internet, Web, Reseña, Recuperación de información.

Title: Web genres: research trends

Abstract: We present a review of some of the main trends of research into web genres. Genres are documents defined primarily by the communicative purpose they accomplish for their producers and recipients, though they also present specific characteristics of topic and form. Thinking of web documents in terms of genres provides a selection criterion which is complementary to topic or subject, and which can also be implemented in automatic classifiers. The review offers a definition of web genres, a summary of achievements in automatic classification of web genres, and a perspective from which to sort out the increasing number of contributions on genres.

Keywords: Web genres, Document types, Linguistic forms, Internet, Web, Review, Information retrieval.

Montesi, Michela. "Géneros web: líneas de investigación". En: *El profesional de la información*, 2008, septiembre-octubre, v. 17, n. 5, pp. 551-557.

DOI: 10.3145/epi.2008.sep.10



Michela Montesi es profesora ayudante en la Universidad Complutense de Madrid. Sus intereses de investigación incluyen la comunicación científica y la teoría del género documental.

1. Introducción

LOS ESTUDIOS SOBRE EL GÉNERO (TIPO O FORMATO DE DOCUMENTO) representan una encrucijada entre varias disciplinas, entre las cuales se sitúan la documentación y la recuperación de información, el análisis del discurso, el diseño de documentos y la interacción hombre-ordenador. Las motivaciones con las que las diversas disciplinas emprenden el estudio de los géneros web son distintas, aunque todas suelen compartir un fundamento teórico común. Se trata básicamente de la idea de que son acciones retóricas tipificadas que se dan en situaciones sociales recurrentes (Miller, 1984).

Dicho de otra forma, para conseguir ciertos objetivos los hablantes nos expresamos por formas

lingüísticas en cierta medida estandarizadas (los géneros), cuyas características varían, además de según el objetivo, por el contexto socio-cultural y la situación en las que se dan. Así en un contexto científico se producen géneros como el artículo de investigación, el resumen o *abstract*, la petición de beca, etc., que desempeñan funciones distintas para sus productores y receptores determinadas por el propio contexto y los diferentes objetivos comunicativos (Swales, 1990).

La teoría de Miller se ha popularizado posteriormente gracias a la aplicación del concepto de género a las dinámicas comunicativas de la gran empresa (Yates; Orlikowski, 1992), es decir a estudios de naturaleza sociológica. El concepto de género, por tanto, nace y se desarrolla en el ámbito de los estudios

literarios y del análisis del discurso. Sus implicaciones, sin embargo, se extienden a otras ramas del conocimiento, en particular a las ciencias sociales, incluida la documentación, como también pone de manifiesto la reciente contribución de Andersen (2008) sobre su aplicación en las ciencias documentales en el *Annual review of information science*.

Para la documentación y la recuperación de información, el interés por la teoría del género estriba en su potencialidad de mejorar la recuperación y la selección de la información, especialmente en colecciones amplias de documentos, como la Web (Kwasnik; Crowston, 2005). Los actuales sistemas de recuperación automática de información tienen una limitación fundamental: saben qué dicen los documentos, pero no qué significan

o para qué objetivos pueden ser útiles (Crowston; Kwasnik, 2003). Conceptualizar la información en términos de género permitiría captar las necesidades informativas de un usuario situado en un contexto determinado y con ciertos objetivos y necesidades informativas. Además, eso también nos permitiría entender cómo los usuarios utilizan la información (Montesi; Navarrete, 2008). Otra de las suposiciones acerca del valor del género para la recuperación de información es que sea posible desarrollar programas capaces de reconocerlos automáticamente. Las implicaciones para la recuperación de información serían notables especialmente en la Web, pero la gestión de información a nivel de organización y empresa se beneficiaría también.

En el resto de este artículo presentamos algunas tendencias o áreas de investigación concernientes a una manifestación concreta de géneros, los géneros web y/o géneros electrónicos, con especial hincapié en aquellas áreas de mayor interés para la documentación. Incluimos tanto los trabajos de clasificadores automáticos de géneros, como aquellos más enfocados en poner de relieve otros aspectos, como su evolución histórica, nacimiento, o funcionalidades.

“La teoría del género puede mejorar la recuperación y la selección de la información, especialmente en colecciones amplias de documentos, como la Web”

1.1. Cuestiones metodológicas

La reseña bibliográfica se limita a estudios realizados expresamente en el marco de la teoría del género. Muchos de los géneros web de los que hablaremos se han estudiado desde perspectivas distintas, pero fuera del marco al que nos referimos. Asimismo, las referencias citadas representan una selección de la bibliografía que se ha producido. El objetivo fundamental de esta contribución es empezar a “poner orden” en un área aún muy joven de la investigación e intentar trazar unas tendencias generales.

Para recopilar las referencias nos hemos basado en varias fuentes. La mayoría de las publicaciones del área aparecen en las actas de la *Hawaii international conference on system science*, que dedica anualmente un minitrack a estos estudios, y la revista *Information processing and management*, donde se van publicando cada vez más contribuciones relevantes. También se realizaron búsquedas en bases de datos especializadas, como *Library and information science abstracts* y *Web of science*, y en buscadores web como *Google academic*, siguiendo asimismo la trayectoria de citas de algunas de las contribuciones más significativas. La palabra *genre* o el nombre del género (*homepage*, *weblog*, etc.) aparece con frecuencia en el propio título. Por otro lado, se trata de un área muy reciente de investigación y abierta a muchas disciplinas, así que los autores pueden asignar descriptores de lo más diverso. La propia expresión géneros web puede sustituirse con *cybergenres* o géneros de internet, etc. Finalmente, se han seguido las publicaciones de algunos individuos activos en un área que sigue siendo bastante reducida.

2. Los géneros web

Con respecto a sus correspondientes *offline*, presentan ciertas particularidades. En primer lugar, se distinguen por las funciones

ofrecidas por el medio en el que se distribuyen, es decir por la posibilidad de búsqueda, el *browsing*, y la interactividad (Shepherd; Waters, 1998, 1999). La interactividad conlleva, además, que estos documentos electrónicos puedan leerse y navegarse al mismo tiempo, según ponen de manifiesto Askehave y Nielsen (2005) que hablan al respecto de bi-dimensionalidad de los géneros web.

La naturaleza cambiante y difícil de capturar que se suele atribuir a todo género tanto impreso como electrónico, en la web se caracteriza en términos de hibridismo e individuación (Santini, 2007). Con esto, se indica que manifestaciones concretas de un género pueden mezclar características de otros, mientras que las idiosincrasias individuales pueden dificultar aún más la tarea de su reconocimiento. Reconocer los géneros inequívocamente es difícil tanto de forma automática como manual. Un estudio con 135 usuarios reales demostraba que a la hora de denominarlos, las opiniones propuestas por los usuarios se disparan (Santini, 2008). Por otro lado, los usuarios parecen conscientes de lo que es un género, incluso en un entorno web, y están de acuerdo acerca de la utilidad del concepto a la hora de utilizar recursos de internet (Eissen; Stein, 2004).

Por otra parte, el problema de si este concepto coincide con el de página web plantea opiniones dispares entre los investigadores. Aunque la mayoría de los estudiosos consideran implícitamente que cada página web es un caso de género, otros extienden el concepto de género a todo el sitio web, mientras que hay quienes miran a componentes de páginas que pueden ocurrir individualmente o en combinación (Haas; Grams, 2000, p. 186). Dong et al. (2008) mencionan la dificultad de reconocer las fronteras entre los géneros, aunque en la práctica adoptan la página web como delimitación natural.

3. Reconocimiento automático de los géneros web

La clasificación de páginas web en base a su género es un fenómeno relativamente reciente. Mientras las colecciones de documentos digitalizados recogían documentos genéricamente iguales (artículos de investigación, resúmenes), la clasificación por tema era suficiente. Con la aparición de conjuntos documentales constituidos por documentos diversos y heterogéneos, como la web, la clasificación por tema se queda corta para capturar los matices en detalle (**Choi; Yao, 2005**). El concepto de género en aplicación al procesamiento del lenguaje natural aparece en un trabajo de **Biber** de 1989, que caracteriza a ciertos textos como géneros según un conjunto de 5 dimensiones lingüísticas recurrentes o escasamente presentes en los mismos. El supuesto de los estudios siguientes, y más recientes, es que es posible desarrollar algoritmos capaces de identificar automáticamente el género de un documento electrónico o web. La información obtenida de esta forma puede emplearse para reducir los resultados de la búsqueda o para establecer prioridades a la hora de presentarlos (**Kwasnik, et al. 2001**).

De hecho, la complementariedad de la clasificación de páginas web en base al género con respecto a la del tema está demostrada por tener en cuenta características y atributos diferentes para las dos modalidades (**Choi; Yao, 2005**). Por ejemplo, las palabras vacías como pronombres, preposiciones y otras palabras comunes, que se eliminan a la hora de representar el contenido de una página en la clasificación por tema, se encuentran incorporados en los clasificadores por género debido a que se considera que no varían entre géneros parecidos.

Los experimentos realizados miden la prestación de diversos clasificadores según los conjuntos

de atributos y características que implementan. Los resultados han sido variables. Según **Kim y Ross (2008)**, la diversidad de los resultados se debe a que géneros diferentes pueden necesitar conjuntos de atributos y características distintos. Los algoritmos desarrollados entonces funcionarían muy bien con algunos y peor con otros, explicando la heterogeneidad de resultados obtenidos.

Entre las contribuciones más recientes podemos mencionar a **Dong, et al., (2008)**, que retoman el trabajo de **Shepherd et al. (2004)**, comparando las prestaciones de diversos conjuntos de atributos genéricos, en concreto los de forma, función, y contenido de los géneros web. Concluyen que, a efectos de la recuperación de información, la combinación de atributos produce siempre mejores resultados que en el caso de que sean individuales, mientras que la precisión y la exhaustividad parecen depender diferentemente del número de características implementadas por cada atributo. La precisión mejora inversamente y la exhaustividad proporcionalmente al número de características unidas algebraicamente. Por otro lado, se ha descartado la hipótesis de que los géneros se puedan reconocer por el vocabulario empleado, pues la combinación de características lingüísticas generales resultó insuficiente para la clasificación de géneros (**Muresan, et al., 2006**). Otros estudios ponen de relieve que los clasificadores automáticos funcionan mejor cuando se tiene en cuenta la relación entre género y tema (**Finn; Kushmerick, 2006; Lee; Myaeng, 2004**). Finalmente, **Lim et al. (2005)** proponen con éxito una combinación de características lingüísticas generales junto con otras específicas de documentos web, extraídas de la url y de las etiquetas html.

4. El contexto de uso

Considerando que el contexto de uso explica los géneros emplea-

dos, sus características, funciones, etc., parece sensato limitar el estudio de los géneros a un entorno más circunscrito. En la Web, por ejemplo, **Rehm (2002)** opta por centrar su investigación en el campo de los servidores web académicos y en los géneros que se dan exclusivamente en el mismo. Otra rama de la investigación persigue elaborar clasificadores automáticos para aplicaciones en ámbitos empresariales. Los criterios de clasificación no cambian sustancialmente con respecto a los géneros web, pero en estos casos es más apropiado hablar de géneros electrónicos debido a la limitación de su distribución.

La finalidad es integrar conocimientos genéricos a los metadatos de gestión de documentos digitales, y así representar mejor los documentos producidos y utilizados en ciertos entornos. En **Karjalainen et al. (2000)**, por ejemplo, se realizó una investigación en una empresa finlandesa consiguiendo un repertorio de 850 géneros en cooperación con los productores y usuarios de los mismos. También destacan los trabajos realizados por **Freund et al. (2005, 2006)** en el contexto de una gran empresa de ingeniería del software, donde también el fin es desarrollar programas capaces de reconocer los géneros en uso en este contexto.

En ámbitos organizativos y empresariales, el estudio de los géneros también se emplea con fines descriptivos de las pautas de comunicación internas de la empresa. **Orlikowski y Yates (1994)**, por ejemplo, analizan los cauces de comunicación inter-organizacional, teniendo en cuenta los repertorios de géneros empleados, los cuales pondrían de manifiesto aspectos del proceso organizativo de la comunidad. **Yates et al. (1999)** estudian la formación de géneros tras la introducción de un sistema de conferencias por ordenador en una firma japonesa.

En su contexto de fruición, es posible ver la mecánica de uso que

subyace. **Østerlund** (2006) discute las combinaciones de géneros (en este caso impresos) en el contexto de un hospital, explicando que pueden adaptarse a circunstancias distintas y desempeñar funciones diferentes. **Montesi y Navarrete** (2008) asocian conjuntos de géneros web a las diversas fases de trabajo de un ingeniero del software. En las etapas iniciales dominan géneros web como los artículos técnicos, las páginas informativas o de productos, las presentaciones..., que proporcionan direcciones o introducción al tema objeto de la búsqueda. En las fases finales, dominan los que son capaces de proporcionar respuestas a preguntas concretas, como por ejemplo los *threads* de grupos de discusiones, o las guías de uso de software.

“En ámbitos organizativos y empresariales, el estudio de los géneros también se emplea con fines descriptivos de las pautas de comunicación internas de la empresa”

5. Perspectiva histórica

Se puede decir que los estudios sobre el género documental combinan con éxito tecnología y sociología, en un sentido amplio que incluya también a los estudios del discurso o las ramas más “sociológicas” de la documentación, como la investigación de conductas informativas. La comunidad reconoce que los avances técnicos para la clasificación automática de los géneros pasan por conocer aspectos como sus propiedades lingüísticas, su historia, sus mecanismos, o las razones de su producción y fruición. De hecho, muchas de las primeras contribuciones sobre los géneros web los tratan desde una

perspectiva histórica. Pueden poner de relieve tanto los procesos de adaptación del medio impreso al electrónico, como las dinámicas de nacimiento y evolución en el entorno de la web.

Crowston y Williams (2000), analizando un conjunto aleatorio de 100 páginas web, las clasifican en dos categorías:

- géneros producto del medio como las *homepages*, y
- otros con claros referentes impresos y con niveles variables de adaptación al medio.

Diferencian además los géneros web con antecedentes impresos adaptados al entorno y los reproducidos sin variaciones del medio impreso. **Ihlström y Henfridsson** (2005), describen longitudinalmente los procesos de digitalización de periódicos escandinavos en el período 1996-2002. En el transcurso de los años, la versión impresa y la digital se han ido configurando como dos géneros complementarios, donde la función de la versión digital sería la de cubrir el intervalo de tiempo entre dos ediciones impresas. **Herring et al.** (2007) miran un conjunto de 450 weblogs recogidos aleatoriamente en el curso del bienio 2003-2004, poniendo de relieve su evolución. En este intervalo, los blogs han evolucionado hacia una mayor textualidad, y los autores (*bloggers*) han ido presentándose cada vez más con sus nombres propios en lugar que con pseudónimos. Otras características en cambio se quedaron más estables, como el número de links o comentarios en las entradas de los blogs, que fueron reduciéndose, demostrando que no son tan interactivos o conversacionales como se cree. El artículo de investigación destacaría entre los géneros adaptados del mundo impreso por su resistencia al cambio (**Mackenzie Owen**, 2007). Sin embargo, el nuevo medio de distribución habría aportado cambios en este género también, enfatizando

el papel de algunos sub-géneros, como las comunicaciones breves, y permitiendo la publicación de material extra online, como apéndices y material técnico (**Montesi; Mackenzie Owen**, 2008).

Otra forma de analizar los géneros web en su trayectoria histórica es compararlos con antecedentes impresos. Es original el análisis que **Hendry y Carlyle** (2006) y **Hendry et al.** (2006) proponen de todas las páginas web basadas en listados de elementos o enlaces, como las de recursos (*resource pages*), o también *hotlists*, *gateways*, portales. Según estos autores, en las páginas de recursos confluyen los principios de dos áreas de investigación tradicionales, es decir la bibliografía y la gestión de colecciones bibliotecarias. Por un lado, una página de recursos es un listado de referencias (éste sería su aspecto bibliográfico). Por otro lado, estas referencias están relacionadas por referirse a un grupo de usuarios común y a temas relacionados (aspecto explicable con principios de la gestión de colecciones bibliotecarias). La transición a la web de la bibliografía supone el paso de la autoría individual, más propia de las bibliografías impresas, a la colectiva.

En la óptica comparativa con antecedentes *offline*, **Herring et al.** (2004) consiguen distinguir varios sub-géneros de weblogs. Los tres tipos principales, tipo-diario, tipo-filtro, y k-blogs (*knoweldge blogs*), representan la adaptación al entorno web de tres géneros distintos: los diarios; los editoriales y cartas al director o editor de periódicos impresos, y de los diarios que los investigadores mantienen de sus proyectos y donde anotan referencias, datos, evolución, etc., acerca de un dominio determinado. Igualmente, otros blogs menos comunes pueden diferenciarse por sus antecedentes *offline*: de viajes (*travel blogs*) tendrían su antecedente escrito en los álbumes de fotos; los *memory blogs* (el autor pone información para uso

posterior) se parecerían a notas de post-it, etc. Visto desde esta óptica, el género blog parece tener una serie de antecedentes y, dependiendo de cuáles sean, es posible diferenciar los sub-géneros.

6. Autoría de los contenidos

Si es verdad que muchos géneros web pueden caracterizarse en términos de sus antecedentes *offline*, está claro que la Web ha dado lugar a nuevos géneros, inconcebibles sin las funciones ofrecidas por el medio. Muchos investigadores apuntan a la *homepage* como ejemplo típico de género puramente web (Dillon; Gushrowski, 2000; Crowston; Williams, 2000). Sin embargo, también somos proclives a indicar como géneros absolutamente nuevos y exclusivos del entorno web aquellos que se caracterizan por una autoría colectiva, como los *wikis* (artículos de *Wikipedia* y parecidos), los *threads* (conversaciones de grupos de discusión), o algunos tipos de páginas de recursos según las describen Hendry y Carlyle. La autoría colectiva obviamente no es ajena a las publicaciones impresas. Incluso podría decirse que un típico trabajo de investigación, por contener referencias a trabajos de otros, también es un trabajo colectivo. Sin embargo, en un estudio de investigación hay un(os) individuo(s) que selecciona(n) las referencias a incluir, y un propósito único moldea el artículo hasta su conclusión. Lo mismo no vale para los géneros que mencionamos aquí, pues algunos están constantemente cambiando por intervención de personas diferentes (*wikis*), mientras que es difícil prever el camino que tomarán otros (*threads*).

Según Emigh y Herring (2005), las diferencias sub-genéricas entre dos enciclopedias disponibles online, *Wikipedia* y *Everything2*, pueden achacarse a las opciones técnicas de los dos sitios, así como

a los mecanismos de control editorial. Debido al control ejercitado por los participantes más activos de la comunidad y por perseguir la homogeneidad estilística de sus contribuciones, *Wikipedia* sería, en su opinión, un claro ejemplo de reproducción de un género impreso en la web. Esto, según Emigh y Herring, choca con el objetivo de *Wikipedia* de incorporar diversas perspectivas en cada artículo. Sin embargo, puede objetarse que el control ejercitado por los miembros es relativo, pues es prácticamente imposible controlar cada artículo en detalle. Luyt et al. (2008) demuestran que cerca del 20% de los errores en el contenido de los artículos se produce en el primer día o en los períodos iniciales, manteniéndose incluso pasadas muchas rondas de revisión. Además esto no disminuye la naturaleza colectiva de la *Wikipedia*.

El género *thread*, que Cox y Morris (2004) definen como género de “preguntas y respuestas”, se da en protocolos de comunicación mediatizada por ordenador como las listas o especialmente los grupos de discusión o forums. El concepto de comunidad de práctica o de interés se suele sacar a colación, pues los participantes en estas conversaciones de pregunta/respuesta comparten un interés común o un área profesional. Cox y Morris (2004) estudian los *threads* producidos tras envíos de correspondencia electrónica a suscritos a una lista de programadores web. Los *threads* que se producen tras cada envío tratan uno de los tres temas: arreglos y soluciones, recomendaciones, o explicaciones sobre cómo hacer algo. Se plantean generalmente preguntas muy determinadas y específicas a las cuales se proporcionan respuestas igualmente concretas. Øystein y Päivärinta (2007) sugieren utilizar el potencial comunicativo de los grupos de discusión para re-definir el género según las exigencias del público al que vaya destinado en un proceso de auto-creación (*auto-*

poiesis). Su experiencia se basa en el fracaso relativo de un sitio creado para promover la e-democracia por algunos ayuntamientos noruegos.

“La Web ha dado lugar a nuevos géneros, inconcebibles sin las funciones ofrecidas por el medio”

7. Otros géneros web

Quizás uno de los géneros más estudiados sea la *homepage* pues, como decíamos, a menudo se la considera el único puro de la web (Dillon; Gushrowski, 2000). Su objetivo es introducir un individuo o una institución al mundo de todos los internautas posiblemente interesados (Roberts, 1998; Saint-Georges, 1998). Shepherd et al. (2004) las clasifican en personales, corporativas, y organizativas, según presenten a un individuo, una empresa, o una organización. Ryan et al. (2003) ponen de manifiesto 4 tipos de *homepage* institucional, y su evolución en un intervalo de 5 años.

Otros géneros menos estudiados son las reseñas (Pollach, 2006), las tesis y disertaciones (Fox et al., 1999), los anuncios (Fortanet, et al. 1999), y el correo basura (Cukier, et al., 2006).

8. Conclusiones: tendencias futuras

El tema de los repertorios genéricos o conjunto de géneros no es nuevo en la literatura sobre los géneros y los géneros web (Yates; Orlikowski, 1994; Bazerman, 1994), y en este texto hemos tenido ocasión de mencionar algunos trabajos que tocan precisamente ese punto. Sin embargo, creemos que su interés está destinado a crecer, puesto

que su potencial es extremadamente rico. Nos permiten describir la conducta informativa de productos y receptores de estos repertorios (Montesi; Navarrete, 2008), y además comprender cómo sus funciones cambian según el contexto (Østerlund, 2006). Particularmente interesante es explicar la naturaleza de los enlaces entre páginas web por los tipos de géneros de las páginas enlazadas. Por ejemplo, Vaughan et al. (2007) encontraron que los enlaces a páginas distintas de la *homepage* de universidades canadienses tendían a ser de naturaleza académica en contraste con los enlaces a la *homepage* que eran a menudo de naturaleza distinta de la académica. Lo contrario, en cambio, sería aplicable a las páginas de empresas, donde los enlaces entre *homepages* tienden a enlazar negocios altamente relacionados (Vaughan et al. 2007a).

Bibliografía

- Andersen, J. "The concept of genre in information studies. Chapter 8". En: *Annual review of information science and technology*, 2008, pp. 339-367.
- Askehave, I.; Nielsen, A. E. "Digital genres: a challenge to traditional genre theory". En: *Information technology and people*, 2005, v. 18, n. 2, pp. 120-141.
- Bazerman, C. "Systems of genres and the enactment of social interactions". En: Freedman, A.; Medway, P. (eds.). *Genre and the new rhetoric*, pp. 79-101. London: Taylor & Francis, 2004.
- Biber, D. "A typology of English texts". En: *Linguistics*, 1989, n. 27, pp. 3-43.
- Cox, A.; Morris, A. "Information dynamics and discourse in a distributed professional community". En: *Proceedings of the 37th Hawaii international conference on system sciences*, 2004.
- Choi, B.; Yao, Z. "Web page classification". En: Chu, Wesley; Lin, Tsau Young (eds.). *Foundations and advances in data mining*. New York: Springer-Verlag, 2005, pp. 221-274.
- Crowston, K.; Williams, M. "Reproduced and emergent genres of communication on the world wide web". En: *The information society*, 2000, v. 16, n. 3, pp. 201-215.
- Cukier, W. L.; Cody, S.; Nesselroth, E. J. "Genres of spam: expectations and deceptions". En: *Proceedings of the 39th Hawaii international conference on system sciences*, 2006.
- Dillon, A.; Gushrowski, B. A. "Genres and the Web: is the personal home page the first uniquely digital genre?". En: *Journal of the American Society for Information Science*, 2000, v. 51, n. 2, pp. 202-205.
- Dong, L.; Watters, C.; Duffy, J.; Shepherd, M. "An examination of genre attributes for web page classification". En: *Proceedings of the 41st Hawaii international conference on system sciences*, 2008.
- Eissen, S. M.; Stein, B. "Genre classification of web pages. User study and feasibility analysis". En: Biundo, S.; Fruhwirth, T.; Palm, G. (eds.). *Advances in artificial intelligence*. Berlin-Heidelberg-New York: Springer, 2004, pp. 256-269.
- Emigh, W.; Herring, S. C. "Collaborative authoring on the web: a genre analysis of online encyclopedias". En: *Proceedings of the 38th Hawaii international conference on system sciences*, 2005.
- Finn, A.; Kushmerick, N. "Learning to classify documents according to genre". En: *Journal of the American Society for Information Science and Technology*, v. 57, n. 11, 2006, pp. 1.506-1.518.
- Fox, E. A.; Macmillan, G.; Eaton, J. L. "The evolving genre of electronic theses and dissertations". En: *Proceedings of the 32nd Hawaii international conference on systems science*, 1999.
- Freund, L.; Clarke, C. L. A.; Toms, E. G. "Towards genre classification for IR in the workplace". En: Ruthven, I. (ed.). *Proceedings of the 1st International conference on information interaction in context*, 2006.
- Freund, L.; Toms, E. G.; Clarke, C. L. A. "Modeling task-genre relationships for IR in the workplace". En: Baeza-Yates, R. A.; Ziviani, N.; Marchionini, G.; Moffat, A.; Tait, J. (eds.). *Proceedings of the 28th Annual international ACM Sigir conference on research and development in information retrieval*, 2005.
- Haas, S. W.; Grams, E. S. "A discussion of four questions arising from a content analysis of web pages". En: *Journal of the American Society for Information Science*, v. 51, n. 2, 2000, pp. 181-192.
- Hendry, D. G.; Carlyle, A. "Hotlist or bibliography? A case of genre on the web". En: *Proceedings of the 39th Hawaii international conference on system sciences*, 2006.
- Hendry, D. G.; Jenkins, J. R.; McCarthy, J. F. "Collaborative bibliography". En: *Information processing and management*, 2006, n. 42, pp. 805-825.
- Herring, S.; Scheidt, L. A.; Bonus, S.; Wright, E. "Bridging the gap: a genre analysis of weblogs". En: *Proceedings of the 37th Annual Hawaii international conference on system sciences*, 2004.
- Herring, S. C.; Scheidt, L. A.; Kouper, I.; Wright, E. "A longitudinal content analysis of weblogs: 2003-2004". En: Tremayne, M. *Blogging, citizenship and the future of media*. London: Routledge, 2007.
- Ihlström, C.; Henfridsson, O. "Online newspapers in Scandinavia. A longitudinal study of genre change and interdependency". En: *Information technology and people*, 2005, v. 18, n. 2, pp. 172-192.
- Karjalainen, A.; Päivärinta, T.; Tyrväinen, P.; Rajala, J. "Genre-based metadata for enterprise document management". En: *Proceedings of the 33rd Hawaii international conference on system sciences*, 2000.
- Kim, Y.; Ross, S. "Examining variations of prominent features in genre classification". En: *Proceedings of the 41st Hawaii international conference on system sciences*, 2008.
- Kwanik, B. H.; Crowston, K. "Introduction to the special issue. Genre of digital documents". En: *Information technology and people*, 2005, v. 18, n. 2, pp. 76-88.
- Kwanik, B. H.; Crowston, K.; Nilan, M.; Roussinov, D. "Identifying document genre to improve web search effectiveness". En: *Bulletin of the American Society for Information Science and Technology*, 2001, December/January, pp. 23-26.
- Lee, Y. B.; Myaeng, S. H. "Automatic identification of text genre and their roles in subject-based categorization". En: *Proceedings of the 37th Hawaii international conference on system sciences*, 2004.
- Lim, C. S.; Lee, C. J.; Kim, G. C. "Multiple sets of features for automatic genre classification of web documents". En: *Information processing and management*, 2005, v. 41, n. 5, pp. 1.263-1.276.
- Luyt, B.; Hsien Aaron, T. C.; Thian, L. H.; Hong, C. K. "Improving Wikipedia's accuracy: is edit age a solution?". En: *Journal of the American Society for Information Science and Technology*, 2008, v. 59, n. 2, pp. 318-330.
- Mackenzie Owen, J. *The scientific article in the age of digitalization*. Amsterdam: Springer, 2007.
- Miller, C. "Genre as social action". En: *Quarterly journal of speech*, 1984, n. 70, pp. 151-167.
- Montesi, M.; Mackenzie Owen, J. "Research journal articles as document genres: exploring their role in knowledge organization". En: *Journal of documentation*, 2008, v. 64, n. 1, pp. 143-167.
- Montesi, M.; Navarrete, T. "Classifying web genres in context: a case study documenting the web genres used by a software engineer". En: *Information processing and management*, 2008, early preview available.
- Muresan, G.; Smith, C. L.; Cole, M.; Liu, L.; Belkin, N. J. "Detecting document genre for personalization of information retrieval". En: *Proceedings of the 39th Hawaii international conference on system sciences*, 2006.
- Orlikowski, W. J.; Yates, J. A. "Genre repertoire: the structuring of communicative practices in organizations". En: *Administrative science quarterly*, 1994, v. 39, n. 4, pp. 541-574.
- Østerlund, C. "Combining genres: how practice matters". En: *Proceedings of the 39th Hawaii international conference on systems science*, 2006.
- Øystein, S.; Päivärinta, T. "Autopoietic cyber-genres for e-democracy? genre analysis of a web-based discussion board". En: *Proceedings of the 38th Hawaii International conference on system sciences*, 2007.

Pollach, I. "Electronic word of mouth: a genre analysis of product reviews on consumer opinion web sites". En: *Proceedings of the 39th Hawaii international conference on systems science*, 2006.

Rehm, G. "Towards automatic web genre identification. A corpus-based approach in the domain of academia by example of the academic's personal homepage". En: *Proceedings of the 35th Hawaii international conference on systems science*, 2002.

Roberts, G. F. "The home page as genre: a narrative approach". En: *Proceedings of the 31st Hawaii international conference on systems science*, 1998.

Ryan, T.; Field, R. H. G.; Olfman, L. "The evolution of US state government home pages from 1997 to 2002". En: *International journal of human-computer studies*, 2003, v. 59, n. 4, pp. 403-430.

Saint-Georges, I. "Click here if you want to know who I am. Deixis in personal homepages". En: *Proceedings of the 31st Hawaii international conference on system sciences*, 1998.

Santini, M. "Characterizing genres of web

pages: genre hybridism and individualization". En: *Proceedings of the 40th Hawaii international conference on system sciences*, 2007.

Santini, M. "Zero, single, or multi? Genre of web pages through the users' perspective". En: *Information processing and management*, 2008, n. 44, pp. 702-737.

Shepherd, M.; Watters, C. "The evolution of cybergenres". En: *Proceedings of the 31st Annual Hawaii international conference on systems sciences*, 1998.

Shepherd, M.; Watters, C. "The functionality attribute of cybergenres". En: *Proceedings of the 32nd Annual Hawaii international conference on systems sciences*, 1999.

Shepherd, M.; Watters, C.; Kennedy, A. "Cybergenre: automatic identification of home pages on the web". En: *Journal of web engineering*, 2004, v. 3, n. 3-4, pp. 236-251.

Swales, J. M. *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press, 1990.

Vaughan, L.; Kipp, M.; Gao, Y. "Are co-linked web sites really related? A link classification stu-

dy". En: *Online information review*, 2007a, v. 31, n. 4, pp. 440-450.

Vaughan, L.; Kipp, M.; Gao, Y. "Why are web-sites co-linked? The case of canadian universities". En: *Scientometrics*, 2007, v. 72, n. 1, pp. 81-92.

Yates, J. A.; Orlikowski, W. J. "Genres of organizational communication: a structural approach to studying communications and media". En: *The academy of management review*, 1992, v. 17, n. 2, pp. 299-326.

Yates, J. A.; Orlikowski, W. J.; Okamura, K. "Explicit and implicit structuring of genres in electronic communication: reinforcement and change of social interaction". En: *Organization science*, 1999, January-December, v. 10, n. 1, pp. 83-103.

Michela Montesi, Facultad de Documentación, Universidad Complutense de Madrid.
montesi@ccdoc.ucm.es

¿Sabías que EPI tiene blog?

Se llama "Zona de notas"

<http://www.elprofesionaldelainformacion.com/notas/>

En él puedes seguir toda la actualidad y noticias sobre "El profesional de la información": novedades, comentarios sobre los contenidos, call for papers, reseñas de actividades, noticias sobre EPI en otros medios, etc.

El profesional de la **información** Indexada por ISI y Scopus

ISSN 1386-6710 

Revista internacional científica y profesional

ARCHIVO DE ZONA DE NOTAS

- INICIO
- ZONA DE NOTAS
- NUMEROS ANTERIORES
- SOBRE EPI
- DISTRIBUCIÓN/BASES DE DATOS
- SUSCRIPCIONES
- PUBLICIDAD
- INFORMACIÓN AUTORES
- AGENDA
- IWETEL
- THINKEPI
- COPYRIGHT
- EQUIPO DE LA REVISTA
- CONTACTO

Curso en la Facultad de Medicina de Salamanca (21-25 de julio de 2008)

Fecha de publicación: 25 Agosto 2008

El curso de verano "Biblioteca virtual en ciencias de la salud. Búsqueda y recuperación de información relevante: un reto para el profesional sanitario e investigador" se va consolidando gracias a la buena fama que ha adquirido con sus ya 3 ediciones.
http://sabus.usal.es/bib_virtual3/

Los asistentes son los que mejor promoción han hecho del mismo con sus comentarios entre los colegas y también en blogs.



Los organizadores: **Helena Martín-Rodero**, jefa de Área Biosanitaria de la Biblioteca de la Universidad de Salamanca, y **Julio Alonso-Arévalo**, jefe de la Biblioteca de la Facultad de Documentación y Traducción.

■ SUSCRIPCIÓN RSS

Ads by Google

Curso Rehabilitación
La guía con la información que buscas para ampliar tu formación.
www.aprendemas.com

ProQuest **¡NUEVO!** Entrepreneurship

**ACCESO SIN PRECEDENTES A UNA AMPLIA SELECCIÓN
HERRAMIENTAS PARA EL ESTUDIO Y LA PRÁCTICA DE LA
INICIATIVA EMPRESARIAL.**

- Videoclips con consejos de emprendedores de éxito (eClipsNet de Cornell University)
- Completas herramientas para emprendedores: plantillas, ejemplos de planes de negocio, casos prácticos (*Business Cases* de Ivey, Darden, Babson) y mucho más.
- Recopilación de investigación de miles de recursos académicos, incluyendo títulos de *Emerald*, *Palgrave* y *SSRN working papers*
- Datos prácticos e investigación de mercado para la evaluación de nuevos productos e ideas, tal como *Snapshots Industry Reports* y *Hoovers Company Profiles*

ProQuest *Entrepreneurship* tiene la misma interfaz de usuario y se puede utilizar al mismo tiempo que otras bases de datos como *ABI/INFORM*, para proporcionar a los usuarios los resultados más completos.



ABI/INFORM®

**LA BASE DE DATOS DE REVISTAS DE ECONOMÍA Y
GESTIÓN LÍDER DEL MERCADO AL ALCANCE DE SUS
MANOS**

- Extensa cobertura de todas las disciplinas en el área de los negocios
- Más de 4.000 títulos, incluido el acceso en exclusiva a *MIT Sloan Management Review*, *Rand Journal of Economics*, *European Management Review*, *The Wall Street Journal*
- Contenido de publicaciones no periódicas como *EIU ViewsWire* Country Reports, y más de 20.000 tesis de economía y negocios en texto completo.
- Líder del mercado en indización de economía y negocios durante más de 30 años.

Para obtener más información o disponer de un acceso en prueba gratuito, envíe un correo electrónico a editor@spain.proquest.com citando la referencia AD 816 08