
El género web como útil al servicio de la recuperación de información

Web genres as tools for selective information retrieval

Michela MONTESI (1) y Blanca GIL URDICIAIN (2)

(1) Facultad de Documentación, c/Santísima Trinidad, 37, 28010 Madrid, montesi@ccdoc.ucm.es

(2) blanca@ccdoc.ucm.es

Resumen

La información acerca de los géneros podría mejorar los procesos de selección y búsqueda de información, especialmente en grandes colecciones de documentos. Se asume que el género revela la intención de uso de un documento en determinada comunidad de usuarios. En esta investigación pretendemos averiguar cuáles son los géneros web más utilizados entre los usuarios, al menos, en el contexto de un grupo de usuarios. Para llevar esto a cabo, definimos un contexto, la Facultad de Ciencias de la Documentación de la Universidad Complutense de Madrid, y un grupo de usuarios específicos de dicho centro, los cuales no se tienen en cuenta normalmente en otras investigaciones sobre la materia. Los usuarios web entrevistados fueron seleccionados entre el personal que estudia y trabaja en dicho centro. Los participantes fueron entrevistados en sesiones de 15 minutos aproximadamente. Hicieron comentarios acerca de las páginas web o los recursos web que utilizan de forma habitual, explicando las razones por las que los encuentran útiles, y su forma de utilizarlos. De forma complementaria, las dos investigadoras clasificaron estas páginas web en géneros web. Obtuvimos un conjunto limitado de géneros web, que se detallan lo suficientemente como para conocer la intención de usos y las mutuas relaciones. Los géneros web representan, hasta cierto punto, algunas de las necesidades de información de los grupos de usuarios considerados. Asimismo, los géneros nos dan una idea de la manera de utilizar la información encontrada en la web. Este conjunto de géneros web puede extrapolarse a otras comunidades de usuarios para contrastar resultados. También puede servir como base para implementar algoritmos que reconozcan exclusivamente conjuntos predeterminados de géneros web.

Palabras clave: Géneros web. Recuperación de información. Comportamiento informacional. Análisis experimental.

1. Introducción

Los estudios sobre el género representan una encrucijada entre varias disciplinas, entre las cuales figuran la documentación y la recuperación de información, y la interacción hombre-ordenador. Los géneros se conciben como ac-

Abstract

Information about genres seems to be able to improve the information search and selection processes, especially from large collections of documents (Crowston y Kwasnik, 2003). It is assumed that genre reveals the intended use of a document within certain user communities. Research has either focused on all the web genres available in the Web, or on single web genres, such as blogs (Herring et al., 2005) or online newspapers (Ihlström and Henfridsson, 2005). However, attempts at studying specific sets of web genres and the ties between them are less common. In particular, in this research it is our intention to find out those that appear to be the most popular among users, or at least among certain groups of users. In order to do so, we define a specific context and user group, which is often left undetermined in much research into genre. The context of our research is a university f, i.e. the Faculty of Documentation of the Complutense University of Madrid. The participants web users are a selection of the people working and studying there. The participants are interviewed in sessions of 15 minutes approximately. They comment on some of the web pages or web resources they routinely use, explain why they find them useful; and how they use them. Complementarily, the two researchers classified these web pages into web genres. The web genres represent to a certain extent some information needs of the user group considered, at least those information needs that they satisfy using web-based resources. and give us an insight into the way users utilize web based information. This set of web genres can be further confirmed and extended to other user communities. It can also serve as a basis for training algorithms that can recognize exclusively predetermined sets of web genres. The implications of these results can improve selective information retrieval.

Keywords: Web genres. Information retrieval. Information behavior. Experimental analysis.

ciones retóricas tipificadas que se dan en situaciones sociales recurrentes (Miller, 1984). En otras palabras, los géneros serían formas lingüísticas en cierta medida estandarizadas, cuyas características varían, además de según el objetivo de los interlocutores, según el contexto

socio-cultural y la situación en los que se dan. De esta forma, en un contexto científico se producen géneros como el artículo de investigación, el resumen, la ponencia, etc., que desempeñan funciones distintas para sus productores y receptores, determinadas por el propio contexto y los objetivos comunicativos del caso (Swales, 1990).

Para la documentación, el interés de 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 y Crowston, 2005). Según los partidarios de la teoría del género, una limitación de los sistemas de recuperación automática de información sería que saben qué dicen los documentos, pero no saben qué significan o para qué objetivos pueden ser útiles (Crowston y Kwasnik, 2003). Conceptualizar la información en términos de género permitiría situar en un contexto determinado las necesidades informativas de un usuario y con ciertos objetivos y necesidades informativas. Además, los estudios sobre los géneros también nos permiten entender cómo los usuarios utilizan la información (Montesi y Navarrete, 2008), pues, como decíamos, responden a objetivos dependientes de la situación. Otras 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 pueden ser notables especialmente en la Web (Finn y Kushmerick, 2006). Asimismo la gestión de información a nivel de organización y empresa podría beneficiarse, como demuestran los trabajos de Freund et al. (2005, 2006) y de Karjalainen et al. (2000) para el desarrollo de sistemas de gestión de documentos de empresa basados justamente en criterios genéricos.

A pesar del optimismo acerca de las potencialidades de la teoría del género aplicada a la recuperación de información, la investigación debe enfrentarse a varios obstáculos. Algunos dependen de la propia naturaleza de los géneros, que es cambiante tanto en el entorno impreso como en el electrónico, otros de las proporciones de la Web. Estos aspectos se caracterizan en términos de *hibridismo* e *individuación* (Santini, 2007), pues manifestaciones concretas de un género pueden mezclar características de otros géneros, mientras que géneros concretos pueden presentar rasgos individuales que obstaculizan su reconocimiento. Con todo esto, individuar a los géneros de forma rotunda es una tarea difícil tanto si se realiza de forma automática como de forma manual. Bien es

verdad, sin embargo, que los usuarios están familiarizados con el concepto de género, incluso en la Web, y reconocen que puede ser de utilidad a la hora de emplear recursos de Internet (Eissen y Stein, 2004).

La investigación ha dado lugar tanto a trabajos interesados en encontrar criterios universales de identificación automática de géneros (Finn y Kushmerick, 2006), como a otros centrados en géneros web individuales, tal como los blogs (Herring et al., 2005) o los periódicos en línea (Ihlström and Henfridsson, 2005). Por otro lado, son menos comunes los intentos de estudiar conjuntos específicos de géneros web y los vínculos existentes entre ellos. Nos planteamos por lo tanto un objetivo intermedio entre encontrar criterios de clasificación universales y centrarnos exclusivamente en un género, que es el siguiente: ¿Cuáles son los géneros web que los usuarios de internet suelen guardar en sus colecciones de favoritos? Nuestra intención es averiguar si hay un patrón genérico en las colecciones de favoritos de un conjunto de usuarios, es decir si los géneros web que suelen guardar tienden a repetirse de una colección a otra.

Los archivos de favoritos son espacios personales de información Web, que consisten en conjuntos informativos altamente relevantes para los usuarios que los crean (Abrams et al., 1998). Se han estudiado en el ámbito de las estrategias de reutilización de información. Parece que, de entre los métodos existentes para este fin, como el historial de navegación o enviar emails de links interesantes a uno mismo, los favoritos son unos de los menos utilizados, pues en un estudio los usuarios los evaluaron como escasos en portabilidad, número de puntos de acceso, capacidad de funcionar como recordatorios, y otros aspectos (Jones et al., 2001). Sin embargo, lo que nos interesa aquí es que se trata de conjuntos de enlaces web significativos para un usuario en concreto, y susceptibles de un análisis en términos de género. El propio usuario y sus actividades, además, representan el punto de unión entre los distintos enlaces.

2. Metodología

Con el objetivo indicado, decidimos analizar un conjunto de datos correspondientes a 8 participantes, 4 estudiantes (3 de quinto curso y uno de doctorado) y 4 administrativos empleados en diversas secciones de la Facultad de Documentación de la Universidad Complutense de Madrid. Consideramos fundamental que los participantes compartieran un contexto común, pues es una premisa para encontrar patrones signifi-

cativos en las colecciones de favoritos que deciden almacenar. El número de participantes que efectivamente han tomado parte en la investigación es superior, pues aún estamos recogiendo datos. Sin embargo, los resultados que presentamos en este trabajo se limitan a 8 participantes.

El material correspondiente a cada participante fueron sus respectivas colecciones de sitios web favoritos por un total de 212 enlaces, y las entrevistas realizadas con cada uno de ellos, de la duración de entre 8 y 14 minutos, aproximadamente. Los participantes se contactaron personalmente, explicándoles objetivos y modalidades de la investigación, así como los términos de su participación. Antes de tomar parte en la investigación, se les entregó un folleto informativo, cuyo fin era informarles sobre la confidencialidad de su participación. Todos tomaron parte en la investigación por libre consentimiento, y las identidades se mantuvieron en el anonimato, así como el contenido específico de los enlaces favoritos. Los datos recogidos se guardarán de forma confidencial hasta que la investigación haya terminado, momento en que serán destruidos.

En cuanto a las entrevistas, se realizaron en el despacho de una de las investigadoras o en el lugar de trabajo de los participantes, según su propia elección. Durante las entrevistas los participantes expresaban primero algunas opiniones generales acerca de la herramienta, y de las modalidades y frecuencia de uso. En segundo lugar, se centraban en unos tres o cuatro de sus favoritos, explicando las circunstancias en las que los habían conocido o encontrado y decidido guardarlos. Enfatizamos especialmente el aspecto circunstancial de su decisión de guardar ciertos enlaces, pues queríamos apreciar todos los posibles factores influyentes en la determinación de la relevancia. El objetivo principal de las entrevistas era proporcionarnos un escenario (el participante, sus actividades, intereses, principios, etc.) que facilitara la tarea de la clasificación. Las entrevistas se grabaron y transcribieron.

Los participantes, terminada la entrevista, nos proporcionaban una copia de sus sitios favoritos. Antes de copiar esos archivos se les dejaba depurar la lista por si querían mantener privados algunos de estos enlaces. Para clasificar las páginas web nos basamos inicialmente en la clasificación desarrollada en Montesi y Navarrete (2008), así como en los criterios de reconocimiento de géneros web indicados ahí. Aunque se trata de una clasificación basada en un solo participante, es la única que conocemos que abarque un conjunto extenso de géneros Web

con sus respectivas descripciones. Esa clasificación, sin embargo, se modificó para adaptarla a la muestra de géneros Web recogida en esta investigación, según explicamos en los resultados.

Los datos se manejaron de forma comparativa entre los estudiantes y los administrativos. Destaca en este sentido que el número total de enlaces favoritos del grupo de estudiantes (150) casi dobla el de los administrativos (62). Quizás esta diferencia pueda atribuirse al hecho de que los administrativos se entrevistaron todos en el lugar de trabajo, mientras que los estudiantes acudieron al lugar de la entrevista con una copia de los archivos de favoritos de sus ordenadores personales, donde pueden grabar más enlaces relativos a ocio y aficiones, por ejemplo. También pueden influir factores relativos al perfil profesional o factores generacionales, aunque no recogimos datos demográficos como garantía de la identidad de los participantes. El número de enlaces favoritos por participante varía entre 3 y 49. Cuatro de los 8 participantes tienen sus favoritos organizados en carpetas según el tema de que traten. De estas colecciones, 3 superan los 35 enlaces y 1 se queda por debajo de este umbral.

En cuanto a la clasificación, los géneros web encontrados se presentan en la sección de resultados. Nos resultó imposible determinar la tipología de algunas de estas páginas (3 no clasificables), mientras que nos fue imposible acceder a otras (11 no visibles) por diversas razones (el link había caducado, al abrirlo daba error, el acceso estaba denegado, etc.).

3. Resultados

3.1. Clasificación de géneros Web

En la Tabla I vemos los resultados de la clasificación de favoritos por género y por grupo. Al final de la comunicación incluimos un glosario con indicaciones acerca de cada uno de los géneros web clasificados en esta investigación.

Tipo de página	Estudiantes		Administ.		Total	
	Nº	%	Nº	%	Nº	%
1. Anuncios	3	2	-	-	3	1.4
2. Artículos y Tutoriales	10	6.7	-	-	10	4.7
3. Blogs	1	0.7	4	6.4	5	2.3
4. Descargas	-	-	2	3.2	2	0.9
5. Guías	1	0.7	1	1.6	2	0.9
6. Homepage total	86	57.3	18	29	104	49
Corporativas	12	8	7	11.3	19	8.9
Organizativas	44	29.3	7	11.3	51	24
Personales	7	4.7	-	-	7	3.3
de Periódicos y						

Revistas	10	6.7	2	3.2	12	5.7
Otras	13	8.7	2	3.2	15	7.0
7. Índices	19	12.7	9	14.5	28	13.2
Archivos de listas	3	2	-	-	5	2.3
8. Informativas	8	5.3	5	8	13	6.1
9. Interactivas	6	4	6	9.7	12	5.7
Catálogos	3	2	-	-	3	1.4
10. Recursos	5	3.3	1	1.6	6	2.8
11. Visuales y multimedia	4	2.7	5	8	9	4.2
12. Otras: contenido	2	1.3	1	1.6	3	1.4
13. Otras: navegación	-	-	1	1.6	1	0.5
No clasificable	-	-	3	4.8	3	1.4
No visibles	5	3.3	6	9.7	11	5.1
Total	150		62		212	

Tabla I. Clasificación de favoritos por género (estudiantes y administrativos)

En cuanto a la distribución de los géneros web en los dos grupos de participantes, las *homepages* representan la mayoría de los enlaces guardados en favoritos, especialmente en el conjunto de estudiantes, mientras que el resto de los géneros representa porcentajes reducidos. Por lo tanto, para poder contrastar mejor los resultados, decidimos agrupar los géneros representados minoritariamente. Para eso, tuvimos en cuenta un estudio relativo a las intenciones de los usuarios cuando utilizan recursos web, y otro centrado específicamente en géneros web. Rose y Levinson (2004) diferencian tres objetivos de búsquedas en Internet: de navegación (para encontrar un sitio conocido), informativos (para aprender algo), y para conseguir ciertos recursos (como descargas, entretenimiento, etc.). En referencia específica a los géneros web, Montesi y Navarrete (2008) diferencian los géneros según la finalidad con la que un usuario los utiliza en: de orientación (como los índices), sin objetivo determinado, introductorios, e instructivos. Teniendo en cuenta estos trabajos y ajustándolos a la muestra objeto de estudio, dividimos los géneros web ahí incluidos en cuatro categorías: a) *homepages* y todas sus variantes, que se clasifican aparte porque llevan elementos navegacionales, de contenido, y transaccionales; b) de navegación, que permiten navegar en la web (índices y recursos); c) de contenido, que presentan contenidos concretos como anuncios, artículos y tutoriales, blogs, informativas, visuales y multimedia; y e) transaccionales, que permiten realizar ciertas operaciones, como descargar un software, acceder a un programa de gestión (sobre todo para los participantes administrativos), o consultar un diccionario en línea.

De esta forma obtuvimos un conjunto de datos más homogéneos. Los gráficos 1 y 2 permiten comparar los resultados entre estudiantes y administrativos:

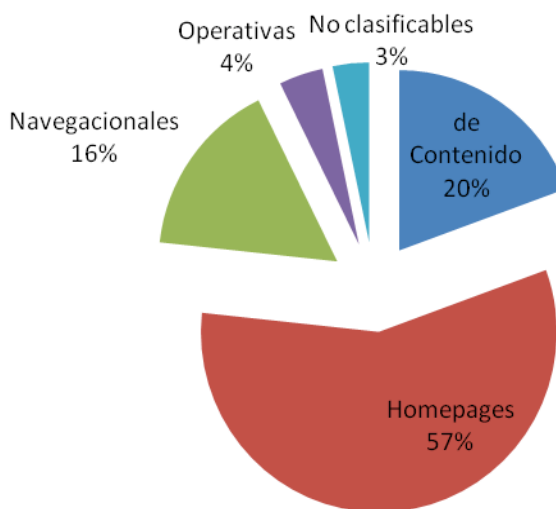


Gráfico 1. Tipos de géneros web – Estudiantes

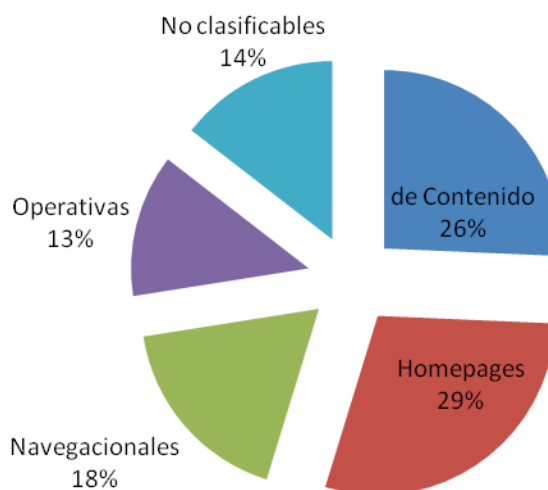


Gráfico 2. Tipos de géneros web – Administrativos

En clave comparativa, destaca la diferencia en la cuota de *homepages* (57% de los favoritos de estudiantes y 29% para los administrativos), y la mayor incidencia de páginas de tipo transaccional para los administrativos (13% frente al 4% de los estudiantes). Es difícil explicar la razón por la que los estudiantes tiendan a guardar más *homepages* que los administrativos. La prevalencia de páginas transaccionales en los administrativos, por otro lado, se debe a que muchas les facilitan acceso a programas de gestión que necesitan para desempeñar sus tareas de trabajo. Considerando estas diferen-

cias, destaca aún más el hecho de que el porcentaje de páginas de navegación y de contenido sea más o menos equivalente en los dos conjuntos de usuarios respectivamente 16% y 20%, y 18% 26% para estudiantes y administrativos.

3.2 Entrevistas

Como decíamos, el objetivo principal de las entrevistas era crear un contexto para la clasificación de los géneros, siendo el objetivo secundario averiguar las circunstancias en las que usuarios de Internet toman decisiones de relevancia. En cuanto a la frecuencia con la que guardan un sitio en favoritos, todos los participantes indicaron que puede variar, pues tienden a guardar muchos sitios en períodos de búsqueda más intensa. Asimismo, manifestaron que sus colecciones van cambiando constantemente, porque tanto añaden como borran enlaces por los que han perdido interés. A continuación resumimos las entrevistas según 3 criterios: cómo encontraron los participantes los enlaces favoritos, por qué los guardaron, y cómo los utilizan.

3.2.1 *Cómo se encuentran los enlaces favoritos*

Los sitios web favoritos se encuentran normalmente utilizando buscadores, incluso para localizar sitios conocidos. Por otro lado, muchos enlaces pueden guardarse en favoritos tras descubrir su existencia en otros entornos que la red: en una clase, leyendo artículos en un periódico impreso; escuchando la radio, porque alguien los recomienda; o por una comunicación oficial de su existencia (por parte de la universidad, por ejemplo). La causalidad también tiene un papel en descubrir los sitios que acaban formando parte de la colección de favoritos. Un participante, por ejemplo, acabó descubriendo la existencia de un traductor utilizando el ordenador de otra persona. Sin embargo, muchas veces pueden encontrarse sitios web interesantes sencillamente buscando; estos sitios pueden ser irrelevantes para la búsqueda de ese momento, pero se consideran interesantes y con un uso futuro.

3.2.2 *¿Por qué se guardan ciertos enlaces en favoritos?*

a) *Porque son únicos.* La razón fundamental para guardar ciertos enlaces es que son *únicos*. El término único recoge un abanico de significados. En primer lugar, la unicidad de los enlaces guardados en favoritos se debe a que buscando información en la web, los sitios que uno encuentra son muchísimos y a menudo de calidad

inferior a las expectativas del participante. Por calidad pueden entenderse cosas distintas: inmediatez, sencillez de uso, y claridad, para algunos; funcionalidades y exhaustividad de la información para otros; posibilidad de personalizar la información y buena estructura, etc. En segundo lugar, unicidad significa rareza no solo en el sentido cualitativo que acabamos de ver. Un participante, por ejemplo, nos comentaba que guardó direcciones de periódicos digitales en una época en la que no había muchos. Un recurso raro, además, es un recurso precioso, pues muchas veces puede haberle costado mucho al participante encontrarlo. Finalmente, el concepto de unicidad conlleva una fuerte carga emocional e ideológica. Se guardan sitios de familiares, por ejemplo, sitios cuya visión política se corresponde con la del participante, que parecen graciosos, o que tratan de la obra literaria preferida. En relación con los factores emocionales, cabe recordar que algunos participantes fueron incapaces de explicar las razones para las que tenían algunos sitios guardados en favoritos, cuando se les preguntaba, por ejemplo, porque no los buscaban directamente en un buscador.

b) *Para ahorrar tiempo.* Un enlace en favoritos permite acceder inmediatamente al recurso que se necesite, traduciéndose en un ahorro de tiempo. El ahorro de tiempo es especialmente significativo cuando el recurso se utiliza a diario incluso por períodos de tiempo limitados (para hacer un trabajo, por ejemplo, o aprender a manejar cierto software). Puede pasar que uno decida añadir un recurso a favoritos después de haberlo visitado varias veces, sobre todo si para alcanzarlo debe pasar por diversas otras páginas.

c) *Para recordar algo.* Como veíamos, muchos participantes conocen la existencia de un sitio web trabajando en otros entornos. Guardar estos enlaces conocidos por otros medios puede ser una estrategia para no olvidarlos. Algo parecido pasa si el participante está chateando o hablando: estando ocupado en otra tarea le resulta útil añadir a favoritos cualquier enlace que sus interlocutores le recomienden, para explorarlo con calma más tarde. Por otro lado, hay direcciones URL especialmente difíciles de recordar: algunas son largas, otras veces, es difícil recordar si el dominio es *.com* o *.es*, etc.

d) *Para realizar un trabajo.* Esta circunstancia solo se da entre los estudiantes, cuando deciden guardar uno o más enlaces que utilizarán con frecuencia para realizar un trabajo.

3.2.3 Cómo se utilizan los sitios favoritos

Los propósitos con los que los usuarios utilizan los enlaces pueden depender del género concreto en cuestión. Sin embargo, no relacionamos expresamente géneros web y propósitos en las entrevistas, así presentaremos motivaciones generales más que relativas a casos concretos. Las principales razones para guardar un enlace son las siguientes: para iniciar la exploración de un sitio web; para comparar y/o tener acceso a varios sitios a la vez (especialmente con páginas de recursos); como recordatorio de sitios buenos o de la trayectoria de búsqueda; para monitorizar posibles cambios o actualizaciones de contenido; para orientación, sobre qué películas ver, sobre qué cocinar, o sobre la planificación de un viaje; para añadir contenido con regularidad (en foros); y finalmente: para realizar algún tipo de transacción (acceder a un programa de gestión, buscar algo, para jugar...).

4. Conclusión

El objetivo del trabajo era encontrar patrones comunes en las colecciones de favoritos, y podemos afirmar que, de acuerdo con los resultados preliminares de la investigación basados en la muestra analizada, esto se ha conseguido. Desde el punto de vista de los estudios sobre el género, hemos visto que estudiar los géneros web en conjuntos (transaccionales, de navegación, etc.) facilita el análisis con respecto a estudios de géneros individuales (índices, blogs, etc.), permitiendo realizar comparaciones y apreciar variaciones. En segundo lugar, la muestra objeto de estudio pone de relieve diferencias entre usuarios en la atribución de relevancia a diversos géneros web, especialmente las *homepages* y las páginas de tipo transaccional. Las *homepages* ocupan más del 50% de los favoritos de los estudiantes, mientras que las páginas interactivas ocupan un espacio porcentual superior en los administrativos. Donde los dos conjuntos de usuarios coinciden es en la relevancia que atribuyen a los géneros de navegación (como los índices y las páginas de recursos) y a los de contenido (como anuncios, artículos y tutoriales, blogs, informativas, visuales y multimedia). En este sentido, la investigación para el reconocimiento automático de géneros web podría indagar las modalidades de diferenciación de estos dos grandes conjuntos de páginas web, de navegación y de contenido, en lugar de centrarse en géneros web específicos o en la web en general.

En referencia a las colecciones de favoritos, esta investigación pone de manifiesto que los archivos de favoritos reflejan la experiencia de

los usuarios con la web, pues incluyen información selecta que afecta a diversas dimensiones existenciales de los participantes (trabajo, estudio, aficiones, familia, etc.). En este sentido sorprende que los trabajos realizados sobre el tema se hayan centrado principalmente en las funcionalidades de la herramienta (Abrams y Backer, 1997; Abrams et al., 1998), más que en su significado en la conducta informativa de los usuarios de Internet, quizás porque sólo es un grupo de usuarios que la utilizan rutinariamente. Las entrevistas, además, nos han permitido apreciar algunos factores que inciden en la determinación de la relevancia, poniendo de manifiesto que la determinación de la relevancia responde a criterios de *unicidad*: del propio participante, de las tareas a desempeñar, del sitio web, y de la situación.

Apéndice: Glosario de géneros web

Artículos y Tutoriales: Los artículos son tanto de tipo periodístico como científicos. Puesto que trabajamos con una muestra reducida de géneros web, nos pareció oportuno incluir los tutoriales en este apartado. Con los artículos comparten el aspecto textual y un propósito instructivo, aunque pueden incluir una cantidad superior de figuras explicativas.

Anuncios: Un aspecto a destacar de estos géneros, cara a su reconocimiento automático, es que siempre proporcionan las coordenadas temporales de algún evento, como, en los casos más frecuentes de nuestra muestra, plazos para solicitar puestos de trabajo o matricularse en ciertos cursos.

Blogs: Puede resultar difícil diferenciarlos de las homepages de tipo personal, aunque la palabra blog, en el texto o en la URL, es decisiva para reconocerlos.

Descargas: Presentan el interfaz de descarga de contenidos o software.

Guías: Las guías son documentos navegables y que se despliegan en varias páginas a las que se puede acceder por medio de un índice general. Su objetivo es instructivo.

Homepages: Siguiendo a Shepherd et al. (2004), las dividimos en personales, corporativas, y organizativas, según describan los intereses de individuos, empresas, o grupos y/o instituciones. También introdujimos una cuarta variante, la de *homepage* de revistas y periódicos, pues ocurría con frecuencia en nuestra muestra. Otras variantes pueden incluir homepages para venta directa al consumidor final o portales.

Índices: Dentro de un sitio, permiten al usuario acceder a contenidos más específicos. Las homepages también cumplen con esta función, pero a su vez cumplen con otras funciones, como presentar la imagen de la empresa o institución. Introdujimos la variante de "archivos de listas" (*mailing list archives*), que entendemos como páginas que introducen, por tema, las conversaciones mantenidas por grupos de discusión.

Informativas: Son páginas de contenido que se centran en describir un elemento del sitio (producto, curso, dependencia de una institución, etc.) o el propio sitio.

Interactivas: Son interfaces que permiten acceder a cierto servicio tras haber insertado datos por parte del usuario, como los interfaces de acceso a cuentas de correo electrónico o de buscadores. En nuestra muestra, destacamos los catálogos de bibliotecas, por pertenecer los participantes a la facultad de documentación.

Recursos: Recogen un conjunto organizado de enlaces externos acerca de un tópico determinado (Hendry et al., 2006). En el conjunto estudiado pueden asumir la forma de directorios, por ejemplo, de productoras cinematográficas. También pueden hospedarse en blogs, cuando es un individuo quien se encarga de recopilar los enlaces.

Visuales y multimedia: Se refieren a toda página cuyo contenido o mensaje se transmita principalmente de forma visual, sirviendo el texto de pura explicación, como mapas, grabaciones, etc. Las páginas visuales están constituidas principalmente por imágenes fijas, como mapas, o por vídeos, como grabaciones de ponencias.

Referencias

- Abrams, D.; Baecker, E. (1997). How people use WWW bookmarks. // CHI 97. (22-27 March 1997) 341-342.
- Abrams, D.; Baecker, E.; Chugnell, M. (1998). Information archiving with bookmarks: Personal web space construction and organization. // CHI 98. Los Angeles CA USA. 41-48.
- Crowston, K.; Kwasnik, B. H. (2003). Can Document-genre metadata improve information access to large digital collections? // Library Trends. 52:2 (2003) 345-361.
- Eissen, S.M.; Stein, B. Genre Classification of Web Pages. User study and feasibility analysis. // Biundo, S.; Fruhwirth, T.; Palm, G. (Eds.). Advances in artificial intelligence, Berlin-Heidelberg-New York: Springer, 2004. 256-269.
- Finn, A.; Kushmerick, N. (2006). Learning to classify documents according to genre. Journal of the American Society for Information Science and Technology. 57:11 (2006) 1506-1518.
- Freund, L.; Clarke, C.L.A.; Toms, E.G. (2006). Towards genre classification for IR in the workplace. // Ruthven, I. (Ed.). Proceedings of the 1st International Conference on Information Interaction in Context, IiX 2006, Copenhagen, Denmark, October 18-20.
- Herring, S.; Scheidt, L.A.; Wright, E.; Bonus, S. (2005). Weblogs as a bridging genre. // Information Technology and People. 18:2 (2005) 142-171.
- Ihlström, C.; Henfridsson, O. (2005). Online newspapers in Scandinavia: A longitudinal study of genre change and interdependency. // Information Technology and People. 18:2 (2005) 172-192.
- Jones, W.; Bruce, H.; Dumais, S. (2001). Keeping found things found on the web, CIKM'01, Noviembre 5-10, Atlanta, Georgia, USA, p. 119-125.
- Karjalainen, A.; Päivärinta, T.; Tyrväinen, P. ; Rajala, J. Genre-based metadata for enterprise document management. // Proceedings of the 33rd Hawaii International Conference on System Sciences, 2000.
- Kwaśnik, B. H.; & Crowston, K. (2005). Introduction to the special issue. Genre of digital documents. // Information Technology and People. 18:2 (2005) 76-88.
- Miller, C. Genre as social action. // Quarterly Journal of Speech. 70 (1984) 151-167.
- Montesi, M.; Navarrete, T. (2008). Classifying web genres in context: a case study documenting the web genres used by a software engineer. // Information Processing & Management. 44:4 (2008) 1410-30.
- Rose, D.E.; Levinson, D. (2004). Understanding user goals in web search. // WWW 2004, May 17-22, New York, New York, USA. 13-19.
- Santini, M. (2007). Characterizing genres of web pages: genre hybridism and individualization. // Proceedings of the 40th Hawaii International Conference on System Sciences, Waikoloa, Big Island, HI, USA, 3-6 January.
- Shepherd, M.; Watters, C. (1999). The functionality attribute of cybergenres. // Proceedings of the 32nd Annual Hawaii International Conference on Systems Sciences, Maui, Hawaii, January 5-8.
- Swales, J.M. Genre Analysis: English in Academic and Research Settings. Cambridge: Cambridge University Press, 1990.