

Usabilidad de los tag-clouds: estudio mediante eye-tracking

Yusef Hassan-Montero
Grupo SCImago, CSIC.

Víctor Herrero-Solana
Departamento de Biblioteconomía y Documentación, Universidad de Granada,

Vicente Guerrero-Bote
Departamento de Información y Comunicación, Universidad de Extremadura,

Resumen

Los tag-clouds, o nubes de etiquetas, son componentes de interfaz en forma de lista compacta de palabras clave, que permiten al usuario explorar y navegar por conjuntos documentales. Si bien en los últimos años han gozado de gran popularidad en el entorno Web, también es cierto que, como interfaces visuales de recuperación de información, presentan evidentes problemas de usabilidad. El presente trabajo se propone indagar en la usabilidad de los tag-clouds, a través de la revisión bibliográfica y un estudio con usuarios utilizando técnicas de eye-tracking o seguimiento visual. Los resultados demuestran que tanto el tamaño de fuente como la forma del tag-cloud tienen una clara influencia en la exploración visual de los usuarios. Respecto a la ordenación de los tags, si bien la ordenación alfabética no ofrece ventajas en tareas de búsqueda exploratoria, los resultados sugieren que la agrupación semántica tampoco supone una mejora en términos de eficiencia en tareas de localización visual de los tags. Finalmente se proponen posibles mejoras en la presentación de los tag-clouds agrupados semánticamente, así como futuras líneas de investigación.

Palabras clave: Tag-clouds, Eye-Tracking, Usabilidad, Recuperación de Información, Búsqueda de Información

Abstract

Tag-clouds are user-interface displays of keywords, which allow users browsing and navigation. Despite their popularity on the Web, tag-clouds present obvious usability problems as information retrieval interfaces. This paper explores usability issues on tag-clouds, by means of bibliographic review and an eye-tracking study. The results show that font size and tag-cloud layout have a clear influence on user visual scanning. Considering that alphabetical ordering does not help users in exploratory search tasks, we would expect semantic clustering to support users in these tasks. However, the results show that semantic clustering does not seem to reduce completion time of visual searching. Finally, design improvements on semantically-clustered tag-clouds are proposed.

Keywords: Tag-Clouds, Eye-Tracking, Usability, Information Retrieval, Information Seeking

1. Introducción

La aparición en la Web de una nueva generación de herramientas sociales, tales como delicious o flickr, trajo consigo la difusión de un nuevo concepto de metadato, los tags o etiquetas. El término tag no siempre se utiliza con el mismo sentido, pero normalmente hace referencia a palabras clave resultado de la indización humana en lenguaje natural. Los tags pueden ser producto de un proceso de indización de autor (como es el caso del etiquetado en flickr o en blogs), o de un proceso de indización social, es decir, cuando son los usuarios o consumidores del recurso web quienes, de forma agregada y distribuida, asignan tags o palabras clave para describirlo (como en el caso de delicious) (Hassan-Montero, 2006).

La popularidad de los tags ha venido acompañada de la de su forma más común de presentación visual, conocida como tag-cloud o nube de etiquetas. Estos tag-clouds son componentes de interfaz que representan conjuntos de tags, normalmente ordenados alfabéticamente, donde la popularidad o frecuencia de cada tag se encuentra codificada mediante atributos visuales (como el tamaño de la fuente), y donde dichos tags actúan de hiperenlaces a los listados de recursos que describen. A diferencia de los listados de enlaces, los tag-clouds se presentan de forma compacta, lo que les confiere ese distintivo aspecto de nube o sopa de palabras (figura 1).

Figura 1: tag-cloud tradicional con ordenación alfabética (Hassan-Montero, Herrero-Solana, 2006).

Aunque la función original de los tag-clouds era la de servir de sistemas de navegación, no parece haber sido precisamente esta función la clave de su éxito, entre otras razones debido a su cuestionable usabilidad. En este sentido, Hearst y Rosner (2007), mediante entrevistas y análisis de ‘webbibliografía’, concluyen que quienes hacen uso de tag-clouds en sus sitios web no lo hacen por su valor como sistema de navegación o visualización, sino motivados principalmente por otras razones: los tag-clouds son percibidos como inherentemente ‘sociales’ o ‘personales’, ya que denotan lo que una persona o grupo de personas están haciendo o en qué están interesados; son visualmente dinámicos, algo que sugiere actividad; resultan una alternativa compacta a las largas listas de enlaces; evidencian que el sitio web usa tags; y por último, los tag-clouds se perciben como algo divertido, popular y de moda.

El presente trabajo se propone indagar sobre la usabilidad de los tag-clouds a través de la revisión bibliográfica y una prueba de evaluación con usuarios mediante la técnica de eye-tracking. El trabajo se estructura de la siguiente forma: El apartado 2 contextualiza el problema de la usabilidad de los tag-clouds en el ámbito de la Recuperación de Información; en el apartado 3 se hace una revisión bibliográfica de los trabajos publicados que ofrecen resultados de evaluación empírica de tag-clouds; en el apartado 4 se describe una prueba de evaluación comparativa, mediante la técnica de eye-

tracking, de un tag-cloud ordenado alfabéticamente y un tag-cloud semánticamente agrupado; y por último, en el apartado 5, se discuten los resultados.

2. Tag-clouds y Recuperación de Información

No cabe duda que los avances científicos y comerciales en el desarrollo de buscadores han facilitado enormemente la tarea de recuperar información en la Web. No obstante, el modelo clásico de buscador en el que el usuario efectúa una consulta y obtiene un listado ordenado de resultados, no ofrece solución a otras necesidades y estrategias de búsqueda, como la búsqueda exploratoria, en la que los usuarios necesitan aprender, descubrir y comprender información de temática desconocida o compleja (Kules, Wilson, Schraefel, Shneiderman, 2008). Para satisfacer un mayor número de necesidades y contextos de búsqueda, adaptarse a la nueva realidad multimedia de contenidos generados cooperativamente y aprovechar la riqueza semántica fruto de la indización social, los buscadores deben inevitablemente evolucionar, no ya sólo desde su dimensión algorítmica, sino también, desde su dimensión visual e interactiva. Como indica Marchionini (2006), con este objetivo, los investigadores están comenzando a combinar las lecciones aprendidas del diseño de interfaces de usuario interactivas con las aprendidas del comportamiento humano en la búsqueda de información, para crear nuevas clases de sistemas de búsqueda que cedan el control de la búsqueda al usuario.

Para ubicar la función que componentes como los tag-clouds pueden tener en un sistema de búsqueda, partimos del marco de búsqueda de información (information-seeking) propuesto por Marchionini y White (2008). Este marco teórico estructura el proceso en diferentes etapas: Primero el usuario reconoce una necesidad de información y acepta el cambio que supone la acción de satisfacerla; la segunda etapa es la actividad de formulación del problema, en el que el usuario conceptualiza los límites de su necesidad e identifica posibles fuentes de información; a continuación el usuario utiliza un sistema de búsqueda para expresar su necesidad de información; en la siguiente etapa el usuario examina el resultado obtenido, resultado que puede llevarlo a reformular y volver a expresar su necesidad; y por último, en algún momento del proceso, el usuario decide finalizar su búsqueda y usar la información.

En este contexto, los tag-clouds podrían asistir al usuario en diferentes etapas del proceso de búsqueda de información. Por un lado, al representar un resumen global de un conjunto documental por medio de los tags o palabras clave más representativos, el tag-cloud puede originar una necesidad en el usuario, asistiendo a aquellos usuarios que en el momento de usar el sistema de búsqueda no tengan definidas claramente sus necesidades. Además, el tag-cloud puede facilitar la expresión de la necesidad en aquellas situaciones en las que el usuario no sea capaz de formular la consulta, pero sí reconocerla entre el conjunto de posibles consultas representadas por cada uno de los tags. En tag-clouds generados dinámicamente a partir de una consulta previa del usuario, es decir, mostrando los tags más representativos asignados a aquellos recursos devueltos por el sistema tras la consulta, estos componentes ofrecerían una guía visual que permitiría al usuario reformular o refinar dicha consulta.

Como vemos, los tag-clouds pueden, potencialmente, servir de complemento en diferentes actividades del proceso de búsqueda de información. Sin embargo, presentan determinados problemas intrínsecos que merman su usabilidad, entendida como el grado con el que el usuario podrá usarlos de forma eficaz, eficiente y satisfactoria. Estos

problemas podemos dividirlos en dos grupos: problemas de usabilidad semántica y problemas de usabilidad visual.

2.1. Problemas de Usabilidad Semántica

En la mayoría de casos, los tags que forman el tag-cloud representan un pequeño conjunto de los que forman la colección total de tags, seleccionados en base a su frecuencia o popularidad. Como es sabido, las palabras más frecuentes son las que menor valor de discriminación poseen (Salton, Wong, Yang, 1975), lo que aumenta la densidad semántica en los tag-clouds, entendida como el grado en que los tags presentes se solapan semánticamente entre sí. En otras palabras, los tag-clouds tienden a estar dominados por unas pocas temáticas diferentes (Begelman, Keller, Smadja; 2006).

No obstante, los problemas de usabilidad semántica pueden mitigarse de forma relativamente sencilla, aplicando para ello funciones de selección de tags basadas en modelos tradicionales de Recuperación de Información. Por ejemplo, en el trabajo de Hassan-Montero y Herrero-Solana (2006) se propone el uso de un método inspirado en la función *tf-idf*, demostrando que reduce el solapamiento entre tags y aumenta su cobertura (porcentaje de recursos que están directamente vinculados a alguno de los tags presentes en el tag-cloud). Como indican los autores, aplicando la función propuesta, desaparecen del tag-cloud numerosos sinónimos o tags vacíos de significado, dando cabida a otros tags con mayor significación y valor de discriminación semántica. Otros trabajos han propuesto el uso de técnicas de SVD (Singular Value Decomposition) para la selección de los mejores candidatos a formar parte del tag-cloud (Provost, 2008), la detección automática de variantes sintácticas mediante funciones de similitud (Astrain, Echarte, Córdoba, Villadangos, 2009) o el uso de algoritmos de ranking (Hotho, Jäschke, Schmitz, Stumme, 2006).

2.2. Problemas de Usabilidad Visual

Los tag-clouds, como representaciones gráficas, son susceptibles de análisis sintáctico-gráfico. La sintaxis gráfica se refiere a cómo los diferentes objetos de la representación se relacionan, ordenan y coordinan entre sí. Por tanto, siguiendo el marco teórico propuesto por Engelhardt (2002), podemos decir que un tag-cloud es un objeto gráfico compuesto por objetos gráficos elementales (tags), que se relacionan visualmente mediante relaciones espaciales y relaciones basadas en atributos visuales.

La relación espacial entre los tags que forman el tag-cloud es una relación por alineación, donde el criterio de ordenación suele ser el alfabético. Este tipo de relación espacial presenta dos problemas de usabilidad: Por un lado, la alineación se encuentra fragmentada, ya que los tag-clouds se presentan en forma de párrafo, exigiendo mayor esfuerzo en la exploración secuencial de los tags (en comparación al esfuerzo que requeriría si se mostraran como una lista simple); por otro lado, el criterio de ordenación alfabética solo resultará útil cuando el usuario se encuentre realizando una búsqueda por ítems conocidos (Rosenfeld, Morville, 2002), es decir, cuando el usuario posea una representación mental sintáctica de su necesidad (Mehlenbacher, Duffy, Palmer, 1989). Como hemos visto al comienzo del apartado 2, las tareas de búsqueda de información en las que parecen ofrecer más ventajas los tag-clouds no son precisamente aquellas en las que el usuario sabe qué está buscando o qué término exacto representa mejor su necesidad de información, ya que de ser así posiblemente encontraría más útil la

consulta directa del buscador. Es por ello que el presente trabajo parte de la premisa de que el criterio de ordenación alfabética en los tag-clouds no ofrece demasiado valor en el contexto de la búsqueda exploratoria.

Además, los tags también se encuentran relacionados mediante atributos visuales, como el tamaño de la fuente, el color o intensidad de dicha fuente y el uso de efectos tipográficos. Como sabemos por la teoría psicológica sobre percepción visual (Ware, 2003; Wolfe, 2006; Wolfe, Horowitz, 2004), propiedades gráficas como el tamaño o el color determinan que nuestra atención visual se vea atraída con mayor fuerza hacia aquellos elementos que presentan mayores diferencias respecto al resto (normalmente los de mayor tamaño o intensidad). Además, estos atributos permiten al usuario detectar de forma automática, a nivel pre-atentivo, relaciones de relevancia o importancia entre elementos (tags). Sin embargo, algunas cuestiones sobre el efecto de estos atributos gráficos en la usabilidad visual de los tag-clouds tienen difícil respuesta desde el análisis teórico, como: ¿Cuál es el efecto del uso combinado de atributos visuales en tag-clouds? ¿Tiene la posición espacial dentro del tag-cloud algún efecto en su capacidad de atracción visual? ¿Hasta qué punto los tags con mayor tamaño o intensidad ocultan la presencia del resto?

2.3. Modelos Visuales Alternativos

A raíz de los problemas de usabilidad descritos, o simplemente fruto de la experimentación, han surgido tanto en la literatura científica como en forma de proyectos personales multitud de modelos visuales alternativos para representar gráficamente conjuntos de tags, algunos muy alejados del concepto original de tag-cloud.

Entre las propuestas presentes en la literatura científica podemos destacar el uso de metáforas topográficas (Fujimura et al., 2008), el geoposicionamiento de tags (Tag-Maps) (Slingsby, Dykes, Wood, Clarke, 2007), el uso del enfoque de los diagramas de Venn para representar grupos de tags y sus relaciones (TagClusters) (Chen, Santamaría, Butz, Therón, 2009), la representación mediante mapas SOM (Self-Organizing Map) (Risi, Mörchen, Ultsch, Lewark, 2007; Lehwerk, Risi, Ultsch, 2008), los ClusterMaps (Klerkx, Duval, 2009), y la presentación en forma de grafos arbóreos (tagFlake) (Caro, Candan, Sapino, 2008) y (TreeCloud) (Gambette, Véronis, 2009). Otros trabajos, como el de Kaser y Lemire (2007), ofrecen diversos algoritmos, tanto para la distribución espacial optimizada en tag-clouds tradicionales, como para la representación de relaciones entre tags.

Uno de los modelos alternativos que mayor impacto ha tenido en la literatura sobre tag-clouds es el de los tag-clouds semánticamente agrupados, propuestos por Hassan-Montero y Herrero-Solana (2006). En este modelo (figura 4) los tags no se encuentran ordenados alfabéticamente, sino agrupados en clusters alineados verticalmente. Dentro de cada cluster los tags se ordenan por alineación horizontal en base a su similitud semántica - similitud calculada sobre la co-ocurrencia relativa entre tags, o coeficiente de Jaccard-. De este modo, se pretende que el usuario pueda detectar e inferir relaciones semánticas entre tags (y clústers de tags), con el objetivo de facilitar tareas de búsqueda exploratoria.



Figura 2: tag-cloud agrupado semánticamente (Hassan-Montero, Herrero-Solana, 2006)

Dado que resulta imposible analizar comparativamente la usabilidad de todos los modelos y variantes de tag-clouds propuestos en la literatura científica, en el presente trabajo nos limitaremos a la evaluación comparativa de la usabilidad visual de los tag-clouds ordenados alfabéticamente y los semánticamente agrupados.

3. Trabajos Relacionados

3.1. Estudios con Usuarios

Uno de los primeros estudios empíricos sobre la usabilidad de los tag-clouds fue el publicado por Kuo et al. (2007). En este trabajo, los autores generan un tag-cloud tradicional a partir de los resultados de búsqueda sobre la base de datos de literatura biomédica PubMed, con el objetivo de comparar la usabilidad de la lista de resultados frente a la del tag-cloud generado. Para ello, se solicitó a los participantes en el estudio que respondieran a una serie de preguntas usando una u otra interfaz. Los resultados indican que con el tag-cloud la calidad de las respuestas es mayor cuando las preguntas son de tipo descriptivo, pero menor que en la lista de resultados en preguntas que requieren identificar relaciones entre diferentes conceptos. Es un resultado lógico si tenemos en cuenta que las primeras se pueden responder únicamente identificando conceptos aislados, mientras que las segundas exigen detectar relaciones que no tienen representación alguna en los tag-clouds tradicionales. Además de la calidad de las respuestas, el estudio mide el tiempo, satisfacción y grado de comprensión de ambas interfaces, métricas en las que la lista de resultados obtiene mejor puntuación que el tag-cloud, menos en el caso del grado de comprensión, donde obtienen la misma puntuación. Entre las conclusiones del trabajo se indica la intención de, en investigaciones futuras, mejorar el tag-cloud mediante agrupación semántica.

Halvey y Keane (2007) estudian comparativamente la usabilidad de los tag-clouds tradicionales y las listas verticales u horizontales de tags, solicitando a los participantes que localicen un tag previamente especificado (búsqueda por elementos conocidos). Los autores detectan que los participantes completan la tarea en menos tiempo mediante las listas simples que con los tag-clouds, un resultado predecible al que ya dábamos explicación en el apartado sobre usabilidad visual, fruto de la fragmentación de la alineación espacial de los tags en los tag-clouds. Además, los autores destacan que los tags con mayor tamaño de fuente son más rápidamente localizables, y que la ordenación alfabética disminuye el tiempo de localización, resultados también predecibles mediante la teoría de la búsqueda guiada (Wolfe, Cave y Franzel, 1989; Wolfe, 2006) y la ley de Hick-Hyman (Hick, 1952; Hyman, 1953; Landauer, Nachbar, 1985) respectivamente.

Una vez más defendemos que el valor de los tag-clouds no parece encontrarse asociado a tareas de localización de ítems conocidos, como la estudiada en el trabajo de Halvey y Keane (2007), sino a tareas de búsqueda exploratoria.

Rivadeneira, Gruen, Muller y Miller (2007) presentan un estudio más exhaustivo, que incluye tanto un marco para la evaluación de tag-clouds como los resultados de dos estudios con usuarios. Según este marco, las tareas que puede soportar un tag-cloud, y por tanto susceptibles de evaluación, son cuatro:

- Búsqueda: Localización de un término específico o uno que represente el concepto deseado.
- Exploración: El tag-cloud se explora sin ningún término o tema específicos en mente.
- Formarse una impresión: Utilizar el tag-cloud como medio para obtener una impresión o comprensión general del conjunto documental subyacente.
- Reconocimiento-Equiparación: Equiparar el conjunto documental representado por el tag-cloud con un conjunto conocido o deseado.

Como se puede observar, menos la búsqueda de términos específicos, el resto de tareas, incluida la búsqueda de términos que representen conceptos deseados, tienen perfecta cabida en la actividad de búsqueda exploratoria.

De las cuatro tareas, la única que es estudiada en el trabajo de Rivadeneira, Gruen, Muller y Miller (2007) es la de “Formarse una impresión”. En el primero de sus experimentos se estudia el efecto que tiene el uso de dos atributos visuales en los tags (tamaño de la fuente y posición espacial) en la capacidad de los participantes para recordar dichos tags tras una exposición breve (20 segundos). Los resultados del experimento muestran que el tamaño de la fuente tiene una clara influencia en el porcentaje de usuarios que recuerdan el tag. Otros resultados obtenidos resultaban menos previsibles, como el hecho de que la proximidad de un tag a otro de gran tamaño no aumenta el porcentaje de participantes que lo recuerdan tras la prueba, o que los tags espacialmente ubicados en el cuadrante superior izquierdo son recordados por más participantes que el resto de tags, lo que sugiere una exploración visual típica de la lectura occidental (izquierda-derecha y arriba-abajo).

En el segundo experimento los autores analizan el efecto que tiene la ordenación de los tags en su capacidad para ser recordados y en las temáticas que se identifican (formarse una impresión). Entre las formas de ordenación se evalúan cuatro: tag-cloud ordenado alfabéticamente, tag-cloud ordenado por la frecuencia de los tags (y por tanto por su tamaño de fuente), tag-cloud desordenado y lista ordenada por frecuencia. Los autores no encuentran influencia entre ordenación y memoria, pero sí una influencia moderada, pero significativa estadísticamente, entre ordenación y capacidad para formarse una impresión. El tipo de ordenación que mejores resultados obtiene es la lista ordenada por frecuencia. Como trabajo futuro planteado por los autores podemos destacar su intención de estudiar estas mismas tareas en tag-clouds agrupados semánticamente.

En el trabajo de Shergold, Good, Gheorghiu y Davies (2007) se estudian las preferencias de los usuarios ante dos opciones de búsqueda: tag-cloud y caja de búsqueda. Para ello, a los participantes se les indica el término de búsqueda (búsqueda por elementos conocidos) y se les presenta una interfaz que incluye tanto caja de

búsqueda como tag-cloud. Los resultados demuestran que la mayoría de participantes decide utilizar el tag-cloud, y que esta opción aumenta en los siguientes intentos, conforme lo participantes se familiarizan con la interfaz. Esto no debe hacernos creer que los tag-clouds se perciben más útiles que las cajas de búsqueda en el contexto de búsquedas por elementos conocidos, ya que seguramente la explicación a este comportamiento se encuentra en que, primero, requiere menos esfuerzo localizar visualmente un tag entre un conjunto reducido de tags que escribir el término en la caja de búsqueda; y segundo, en el experimento el término de búsqueda siempre se encontraba presente entre los que formaban el tag-cloud, lo que provoca precisamente que en los siguientes intentos aumente el uso del tag-cloud frente a la caja de búsqueda. En nuestra opinión, aumentar simplemente el número de tags del tag-cloud sería suficiente para ver un cambio en el patrón de elección, ya que entonces la tarea de exploración visual resultaría más costosa que la de interrogar al buscador. Un dato interesante que ofrece este mismo estudio es que la mayoría de los participantes no son conscientes de que los tags se encuentran ordenados alfabéticamente, un hecho que reforzaría la premisa de que es una ordenación que ofrece muy poco valor en tareas de búsqueda de información.

Sinclair y Cardew-Hall (2008) llevan a cabo un estudio con propósitos similares al anterior, pero donde los participantes llevan a cabo un mayor número y diversidad de tareas. Los resultados muestran que una mayor parte de las tareas encomendadas a los participantes fueron resueltas usando el tag-cloud, pero que en el cuestionario final los participantes afirmaron su preferencia por la caja de búsqueda para 6 de las 10 tareas encomendadas. Las tareas en las que los usuarios parecen mostrar su preferencia por el uso del tag-cloud son aquellas de propósito general, o cuando el término específico buscado se encuentra presente en el tag-cloud. Como concluyen los autores, los resultados sugieren que los tag-clouds son útiles en búsquedas exploratorias, para necesidades no definidas claramente, en el descubrimiento de información por serendipia, o como herramienta de resumen visual, ayudando al usuario a formarse una impresión del conjunto documental.

Otro estudio significativo es el realizado por Bateman, Gutwin y Nacenta (2008), donde los autores analizan la relación entre 9 atributos gráficos diferentes y su influencia visual (capacidad de atraer la atención del usuario). Los atributos gráficos estudiados son: tamaño de la fuente, grosor de la fuente (negrita), color (matiz), intensidad del color, número de pixels (hay determinadas letras que ocupan más pixels que otras), ancho del tag, número de caracteres del tag, área ocupada por el tag y posición espacial en el tag-cloud. Los resultados del estudio señalan que determinados atributos gráficos tienen un claro impacto en la influencia visual (tamaño de la fuente, grosor de la fuente, color y saturación), mientras que el resto tiene un impacto mucho menor. En el caso de la posición espacial, aunque la influencia no es demasiado significativa, los resultados sugieren que las posiciones centrales podrían tener mayor influencia que las superiores e inferiores. El estudio también concluye que, cuando son usados conjuntamente varios atributos gráficos diferentes, ninguno de ellos destaca sobre el resto.

El primer trabajo que ha evaluado de forma comparativa la usabilidad visual de los tag-clouds tradicionales y los tag-clouds semánticamente agrupados ha sido el publicado por Schrammel, Leitner y Tscheligi (2009). En su estudio los autores se proponen analizar el impacto que la ordenación espacial de estos tag-clouds tiene en el tiempo de búsqueda, satisfacción de uso y la memoria, tanto en tareas de búsqueda por elementos

conocidos como en búsquedas de propósito más general. En el primer tipo de tareas, a los participantes se les indicaba explícitamente el término a buscar, mientras que en el segundo caso a los participantes se les solicitaba que buscasen un tag semánticamente relacionado con el que se les presentaba en pantalla.

Los resultados sugieren que para tareas de búsqueda por elementos conocidos, el tag-cloud alfabético ofrece mejores resultados que los tag-clouds semánticamente agrupados en términos de tiempo necesario para completar la búsqueda. Sin embargo, en la prueba de búsqueda de propósito más general, no se aprecian diferencias. Tampoco se encuentran evidencias de que el tipo de ordenación de los tag-clouds influya en la capacidad para recordar tags.

El hecho de que los tag-clouds semánticamente agrupados no parezcan ofrecer ventaja sobre los alfabéticos en tareas de búsqueda generales, es un resultado no esperado por los autores, quienes lo achacan a que si la agrupación no es de suficiente calidad, es posible que los usuarios no sean capaces de diferenciarla de una ordenación aleatoria. No obstante, como los autores señalan, alguna ventaja debe poseer el tag-cloud agrupado semánticamente cuando los participantes lo prefieren y terminan sus búsquedas en menos tiempo que frente a un tag-cloud con los tags ordenados de forma aleatoria.

3.2. Estudios mediante Eye-Tracking

El concepto de eye-tracking hace referencia a un conjunto de tecnologías (hardware y software) que permiten monitorizar y registrar la forma en la que una persona mira una determinada escena o imagen, en concreto en qué áreas fija su atención, durante cuánto tiempo y qué orden sigue en su exploración visual (Hassan-Montero, Herrero-Solana, 2007). El funcionamiento de estos sistemas se basa en el uso de cámaras (eye-trackers) que proyectan rayos infrarrojos hacia uno o los dos ojos del participante, infiriendo la zona de la interfaz que está siendo atendida visualmente en cada momento.

Esta tecnología ofrece información exclusiva y de gran valor para comprender el comportamiento de los usuarios y así poder analizar, con mayor detalle, la usabilidad visual interfaces gráficas. De hecho, han sido numerosos los autores que han señalado la necesidad de llevar a cabo estudios mediante eye-tracking para poder obtener una mayor comprensión acerca de la usabilidad de los tag-clouds (Hearst, Rosner, 2007; Rivadeneira, Gruen, Muller y Miller, 2007; Shergold, Good, Gheorghiu, Davies, 2007; Bateman, Gutwin y Nacenta, 2008; Halvey, Keane, 2008; Schrammel, Leitner, Tscheligi, 2009).

El único trabajo que, hasta donde conocemos, ha evaluado la usabilidad de los tag-clouds mediante eye-tracking, ha sido el publicado por Schrammel, Deutsch y Tscheligi (2009). Los resultados de dicho estudio muestran que los participantes, en su mayoría, reconocen el principio de ordenación alfabética en tag-clouds. No obstante el patrón de exploración visual más utilizado es el aleatorio, y en algunos casos el de escaneado en forma de zig-zag. Además, los resultados muestran que la influencia visual del tamaño de la fuente no varía en función de si el tag-cloud se encuentra ordenado alfabéticamente o por agrupación semántica. Por último, los resultados también confirman los obtenidos por Rivadeneira, Gruen, Muller y Miller (2007), sobre la

influencia visual de la posición del tag, ya que los participantes pasan más tiempo atendiendo al cuadrante superior-izquierdo que al resto.

4. Experimento: Prueba de Usuarios mediante Eye-Tracking

4.1. Materiales y Participantes

Para llevar a cabo el experimento se utilizó el sistema de eye-tracking Tobii T60 y el software de análisis Tobii Studio enterprise edition. El modelo T60 tiene la apariencia de un monitor TFT de 17 pulgadas de color negro, y las cámaras (eye-trackers) se encuentran camufladas en la zona inferior, resultando invisibles a los ojos del participante. Entre las características técnicas del sistema, señalar que usa una frecuencia de 60Hz, tiene una resolución máxima de 1280x1024 pixels, y una precisión de 0.5 grados con un desvío inferior a los 0.3 grados.

El experimento se llevó a cabo en las fechas comprendidas entre el 13 de Marzo y el 15 de Abril de 2008. Fueron reclutados 17 participantes, 6 mujeres y 11 hombres con edades comprendidas entre los 25 y los 49 años. De estos participantes, 8 tenían el título de licenciatura o ingeniería, 4 el de master y 5 el título de doctor; 8 participantes pertenecían al área de la Biblioteconomía y Documentación, otros 8 a Comunicación Audiovisual, y 1 a la Informática.

4.2. Procedimiento

El procedimiento utilizado en el experimento es muy similar al descrito por Schrammel, Leitner y Tscheligi (2009) y Schrammel, Deutsch, Tscheligi (2009). Lo que se pretende es comparar la eficacia y eficiencia con la que los usuarios llevan a cabo búsquedas semánticas y sintácticas sobre un tag-cloud alfabético y uno agrupado semánticamente. Con este objetivo, se definieron cuatro tareas de búsqueda, dos semánticas y dos sintácticas:

- Búsquedas semánticas:
 - BSEM1: Buscar y hacer clic en un concepto relacionado con el Cristianismo (objetivo esperado: “religion”)
 - BSEM2: Buscar y hacer clic en un concepto relacionado con los dibujos y tebeos (objetivo esperado: “comics”)
- Búsquedas sintácticas:
 - BSIN1: Buscar y hacer clic en el nombre del sistema operativo más conocido del mercado (objetivo esperado: “windows”).
 - BSIN2: Buscar y hacer clic en el nombre del buscador más famoso de Internet (objetivo esperado: “google”)

Como se puede apreciar, en las búsquedas sintácticas no se les indica a los participantes específicamente el término a buscar, sino que se intenta naturalizar la tarea, para que sea cada participante quien identifique mentalmente el término buscado.

Los participantes se dividieron en dos grupos. Todos llevaron a cabo las cuatro tareas, una semántica y una sintáctica sobre el tag-cloud alfabético, y las otras dos (semántica y sintáctica) sobre el tag-cloud agrupado semánticamente. La diferencia entre un grupo y otro es que aquellas tareas que un grupo realizaba sobre el tag-cloud alfabético, los

participantes del otro grupo las realizaban sobre el tag-cloud agrupado semánticamente, y viceversa.

4.3. Resultados

4.3.1. Tiempo

En la figura 3 se representa el tiempo promedio que los participantes tardaron en completar cada tarea, tanto en el tag-cloud ordenado alfabéticamente como en el tag-cloud agrupado semánticamente. Estos primeros resultados no parecen confirmar la hipótesis de que el tag-cloud agrupado semánticamente suponga una mejora frente al alfabético para búsquedas semánticas, al menos no en términos de eficiencia de búsqueda. Por otro lado, la única tarea en la que el tag-cloud agrupado semánticamente ofrece mejores resultados que el alfabético, es una tarea de tipo sintáctico (BSIN1), lo que tampoco deja claro que los participantes hayan encontrado ventaja en la ordenación alfabética, ya sea porque no hayan reconocido dicha ordenación en todos los casos, o porque optaron por estrategias de búsqueda aleatoria.

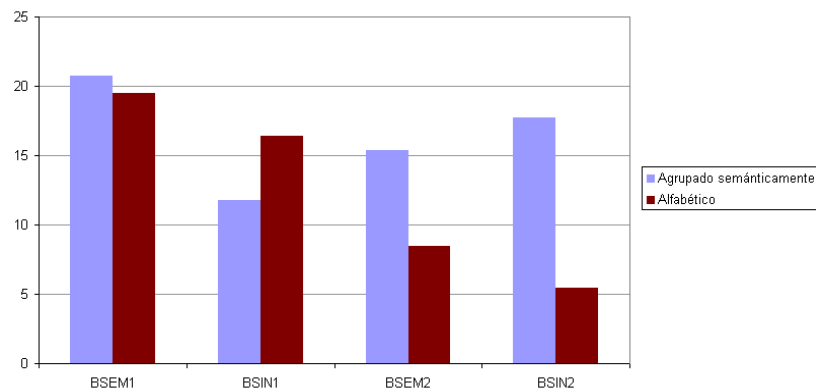


Figura 3: Tiempo promedio empleado para completar las tareas

La figura 4, que representa el tiempo promedio transcurrido hasta que se produce la primera fijación visual en el área ocupada por el tag, corrobora los resultados. La diferencia entre estos tiempos y los representados en la figura 3 es el tiempo destinado a la identificación mental del tag y a la tarea motora de localización y selección mediante el puntero del ratón.

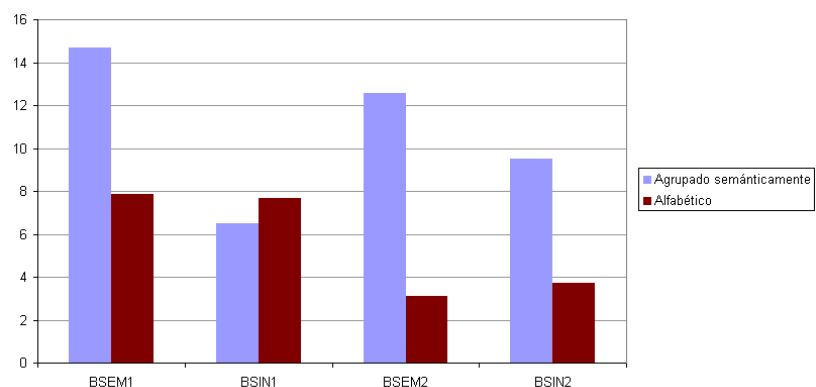


Figura 4: Tiempo promedio hasta que se produce la primera fijación visual en el término objetivo.

4.3.2. Atención Visual

En este apartado se analizan los patrones de exploración visual de los participantes a través de su representación agregada en forma de *HeatMaps* o mapas de calor. En estas representaciones, las zonas de color más 'caliente' (rojo) o de mayor intensidad indican dónde los participantes han realizado mayor número de fijaciones visuales, y los iconos en forma de ratón, dónde han hecho clic.

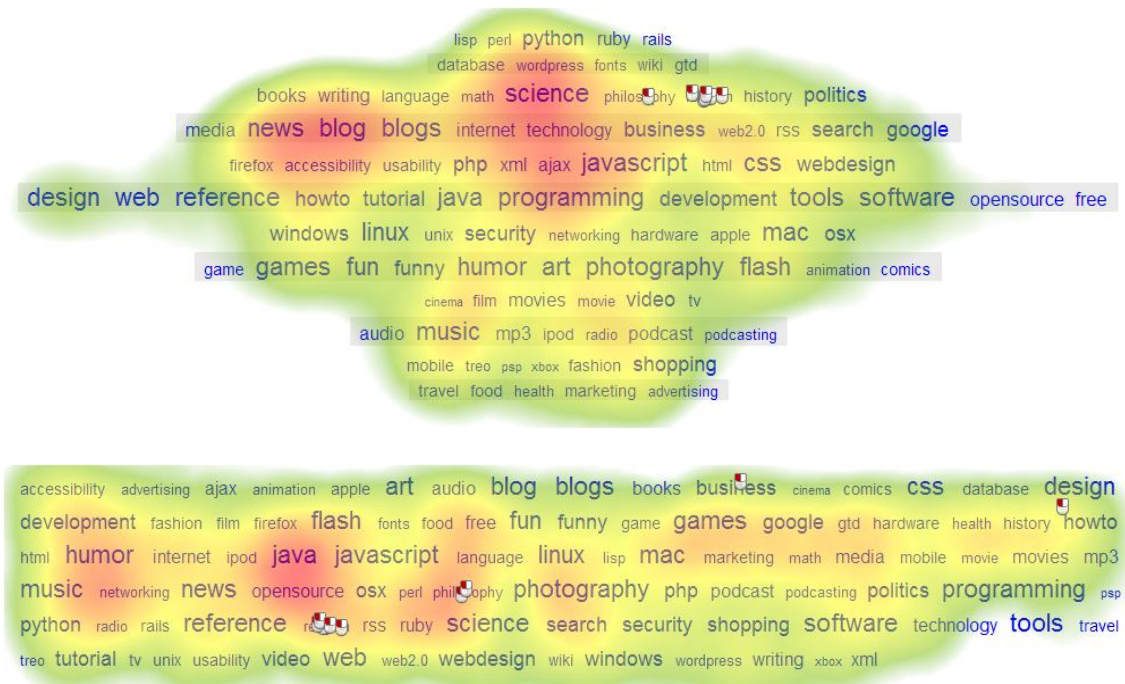


Figura 5: Mapas de calor para la tarea BSEM1

Lo primero que se puede confirmar a través de los mapas de calor es que los tags de mayor tamaño reciben mayor número de fijaciones visuales, corroborándose los resultados de estudios anteriores sobre la influencia visual del tamaño de fuente en tag-clouds.

Aunque una gran parte de los participantes utilizaron estrategias de búsqueda aleatoria, tanto en el tag-cloud agrupado semánticamente como en el alfabético, los mapas de calor indican que la forma visual de los tag-clouds influye en los patrones de exploración. Por ejemplo, en el tag-cloud alfabético se aprecia un patrón de exploración en forma de F (Nielsen, 2006), probablemente motivado por el aspecto de párrafo del tag-cloud y por el orden alfabético de los tags. Esto confirma los resultados de aquellos trabajos previos (Rivadeneira, Gruen, Muller, Miller, 2007; Schrammel, Deutsch, Tscheligi, 2009) que afirman que el cuadrante superior del tag-cloud presenta más influencia visual que el resto de zonas, aunque sólo para el caso de los tag-clouds alfabéticos. Como se puede observar analizando los *HeatMaps*, en el tag-cloud agrupado semánticamente esta influencia no resulta evidente.

En el caso del tag-cloud agrupado semánticamente, además, hay menos zonas que pasen desapercibidas o queden desatendidas visualmente, aunque esto puede ser simplemente consecuencia del mayor tiempo empleado por los participantes para completar las tareas en este tag-cloud. Sin embargo, resulta destacable que determinados tags en el tag-cloud

agrupado semánticamente hayan recibido muchas más fijaciones que en el tag-cloud alfabético, como es el caso de ‘Science’, lo que sugiere, de nuevo, que la forma visual del tag-cloud y la posición de los tags serían factores de influencia visual.

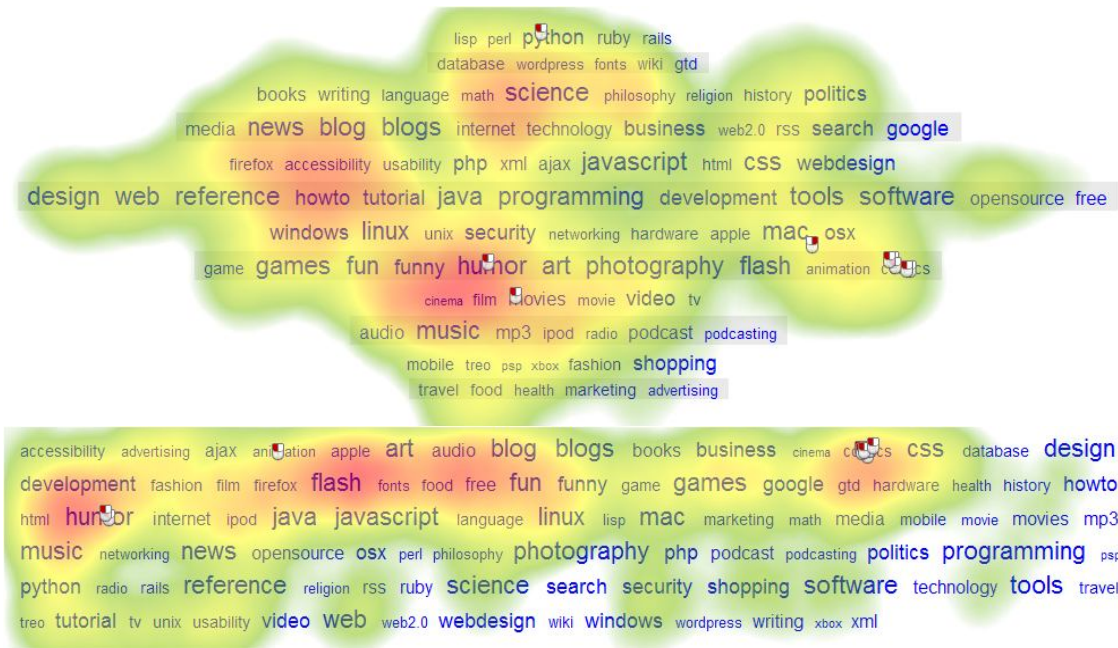


Figura 6: Mapas de calor para la tarea BSEM2

Otra diferencia destacable es que en el tag-cloud alfabético, los movimientos ‘sacádicos’ (saltos rápidos del ojo de un punto de fijación a otro), se producen más frecuentemente de forma horizontal o en zig-zag, mientras que en tag-cloud agrupado semánticamente los movimientos suelen seguir un patrón más aleatorio. Esto evidencia que la ordenación mediante agrupación semántica no es reconocida explícitamente por los usuarios, ni por tanto aprovechada para la consecución de las tareas.

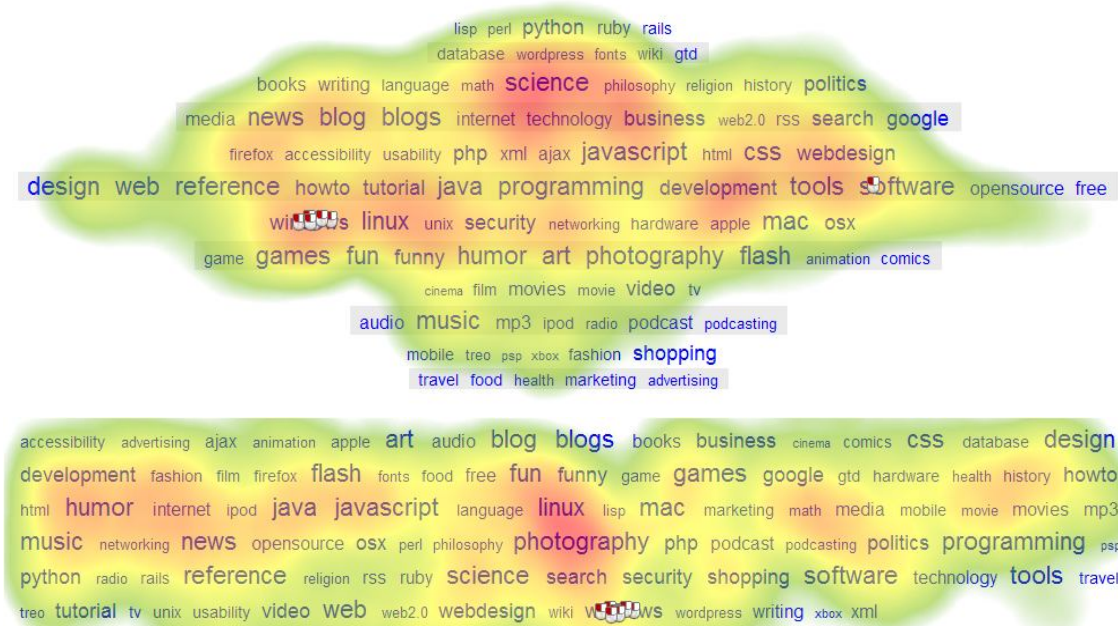


Figura 7: Mapas de calor para la tarea BSIN1

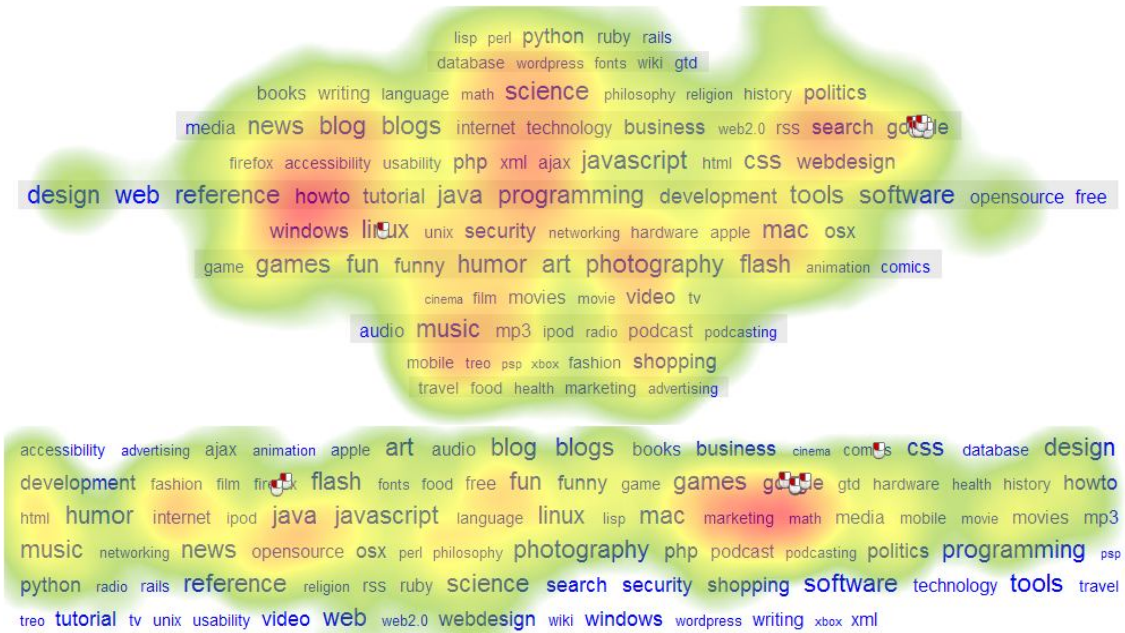


Figura 8: Mapas de calor para la tarea BSIN2

Como se puede observar en los *HeatMaps*, varios participantes en diferentes tareas terminaron seleccionando un tag no esperado o predecible. La mayoría de casos se deben a errores de comprensión de la tarea o falta de conocimiento para ejecutarla correctamente, sin embargo, otros casos son reflejo de la naturaleza multidimensional de las relaciones semánticas en el conocimiento humano. Por ejemplo, en la tarea BSEM1, en la que se debía buscar un tag relacionado con el Cristianismo (como es ‘religion’), hubo un participante que, comprendiendo perfectamente la tarea, seleccionó ‘business’.

5. Discusión

Como se argumenta en este y otros trabajos (Sinclair y Cardew-Hall, 2008), los tag-clouds pueden resultar útiles en tareas de búsqueda exploratoria, en las que los usuarios buscan formarse una impresión acerca del conjunto documental representado, descubrir información desconocida, y reconocer y formarse necesidades de información.

La ordenación alfabética, apropiada para búsquedas de elementos conocidos, en el contexto de búsqueda exploratoria no parece ofrecer demasiado valor para el usuario. Además, como señalan Halvey y Keane (2007), las listas simples ordenadas alfabéticamente incluso ofrecen mejores resultados para este tipo de tareas que los tag-clouds.

En el experimento descrito en el presente trabajo, así como en otros publicados previamente (Halvey, Keane, 2007; Rivadeneria, Gruen, Muller, Miller, 2007; Bateman, Gutwin y Nacenta, 2008; Schrammel, Deusch, Tscheligi, 2009), se demuestra la influencia visual que tienen atributos gráficos como el tamaño de la fuente de los tags. Probablemente el fuerte peso de esta influencia visual es la que provoca patrones de exploración “aleatorios”, es decir, que los usuarios no siempre suelen reconocer, ni por tanto aprovechar en sus búsquedas, criterios de ordenación en los tag-clouds. También es cierto que posiblemente, en tag-clouds de mayor tamaño que los estudiados hasta la

fecha en la literatura científica, los usuarios se verían abocados a identificar algún tipo de ordenación, pues la exploración visual aleatoria dejaría de ser percibida como una estrategia eficiente de búsqueda.

Aunque Rivadeneria, Gruen, Muller y Miller (2007) y Schrammel, Deusch y Tscheligi (2009) detectan que el cuadrante superior-izquierdo del tag-cloud presenta una mayor influencia visual, los resultados Bateman, Gutwin y Nacenta (2008) indican que son las zonas centrales las de mayor influencia. Los resultados del presente trabajo sugieren que la forma visual del tag-cloud puede influir en la exploración visual, condicionando la influencia que determinadas posiciones tienen en la probabilidad de ser foco de atención, lo que explicaría las diferencias obtenidas en los resultados de los trabajos mencionados.

El presente trabajo, junto a los realizados por Schrammel, Leitner y Tscheligi (2009) y Schrammel, Deusch y Tscheligi (2009), son los únicos que han abordado hasta la fecha la comparación empírica de tag-clouds alfabéticos y agrupados semánticamente. Los resultados de estos estudios no demuestran ventaja de los tag-clouds agrupados semánticamente frente a los alfabéticos para búsquedas semánticas, al menos no en términos de eficiencia o tiempo necesario para completar las tareas.

El principal problema de los tag-clouds agrupados semánticamente es que los usuarios no reconocen la ordenación semántica. Schrammel, Leitner y Tscheligi (2009) lo achacan a la posible falta de calidad de la agrupación semántica del tag-cloud utilizada en su estudio, y Schrammel, Deusch y Tscheligi (2009) afirman que no reconocer esta ordenación no parece ser la única causa de que los usuarios no la aprovechen. En el presente trabajo creemos que la razón principal por la que los usuarios no reconocen ni aprovechan esta agrupación semántica hay que buscarla en su diseño visual, en concreto en la forma de representar gráficamente estas relaciones de agrupación. Entre las posibles mejoras creemos que se podrían ordenar los clusters como pilas verticales en vez de horizontales, así como utilizar tags representativos para cada cluster - extraídos mediante técnicas automáticas - a modo de rótulos descriptivos.

Otro aspecto discutible acerca de los resultados analizados, se encuentra en relación con el enfoque de evaluación empleado tanto en el presente trabajo como en estudios previos. Este es un enfoque de evaluación reduccionista, basado en experimentos controlados en laboratorio y que emplea como unidad de análisis tareas básicas predefinidas. Sería recomendable, en posteriores estudios, indagar en el uso de metodologías de evaluación a largo plazo (Shneiderman, Plaisant, 2006), que permitan analizar la usabilidad de estas interfaces en espacios de tiempo prolongados y contextos naturales de trabajo.

Agradecimientos

A la Asociación Interacción Persona-Ordenador (AIPO), y las empresas Alt64 y Tobii Technology por la concesión, en calidad de préstamo, del hardware y software necesarios para el experimento mediante la técnica de eye-tracking.

Referencias

Astrain, J.J.; Echarte, F.; Córdoba, A.; Villadangos, J. (2009). A Tag Clustering Method to Deal with Syntactic Variations on Collaborative Social Networks. En: Lecture Notes in Computer Science, Web Engineering, Vol. 5648/2009, pp. 434-441.

Bateman, S.; Gutwin, C.; Nacenta, M. (2008). Seeing things in the clouds: the effect of visual features on tag cloud selections. Proceedings of the nineteenth ACM conference on Hypertext and hypermedia, Pittsburgh, PA, USA, pp. 193-202.

Begelman, G.; Keller, P. and Smadja, F. (2006). Automated Tag Clustering: Improving search and exploration in the tag space. WWW2006, May 22–26, 2006, Edinburgh, UK.

Caro, L.D.; Candan, K.S.; Sapino, M.L. (2008). Using tagflake for condensing navigable tag hierarchies from tag clouds. Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining, Las Vegas, Nevada, USA, pp. 1069-1072.

Chen, Y.X.; Santamaría, R.; Butz, A.; Therón, R. (2009). TagClusters: Semantic Aggregation of Collaborative Tags beyond TagClouds. Lecture Notes in Computer Science, Springer, Vol. 5531/2009, pp. 56-67.

Engelhardt, J. (2002). The Language of Graphics: A framework for the analysis of syntax and meaning in maps, charts and diagrams. Tesis Doctoral, Institute for Logic, Language and Computation. Universidad de Amsterdam, 2002. ISBN 90-5776-089-4. Disponible en: <http://www.yuriweb.com/>

Fujimura, K.; Fujimura, S.; Matsubayashi, T.; Yamada, T.; Okuda, H. (2008). Topigraphy: Visualization for Large-scale Tag Clouds. WWW2008, April 21–25, 2008, Beijing, China, pp. 1087-1088.

Gambette, P.; Véronis, J. (2009). Visualising a Text with a Tree Cloud. International Federation of Classification Societies 2009 Conference. Disponible en: <http://www.lirmm.fr/~gambette/2009GambetteVeronis.pdf>

Halvey, M.J.; Keane, M.T. (2007). An assessment of tag presentation techniques. Proceedings of the 16th international conference on World Wide Web, Banff, Alberta, Canada, pp. 1313-1314.

Hassan-Montero, Y. (2006). Indización Social y Recuperación de Información. En: No Solo Usabilidad, nº 5, 2006. ISSN 1886-8592. Disponible en: http://www.nosolousabilidad.com/articulos/indizacion_social.htm

Hassan-Montero, Y.; Herrero-Solana, V. (2006). Improving Tag-Clouds as Visual Information Retrieval Interfaces. InSciT 2006: Mérida, Spain. October 25-28, 2006. Disponible en: http://www.nosolousabilidad.com/hassan/improving_tagclouds.pdf

Hassan-Montero, Y.; Herrero-Solana, V. (2007). Eye-Tracking en Interacción Persona-Ordenador. En: No Solo Usabilidad, nº 6, 2007, ISSN 1886-8592. Disponible en: <http://www.nosolousabilidad.com/articulos/eye-tracking.htm>

Hearst, M.A.; Rosner, D. (2008). Tag Clouds: Data Analysis Tool or Social Signaller?. Proceedings of the Proceedings of the 41st Annual Hawaii International Conference on System Sciences, pp. 160.

Hick, W.E. (1952). On the rate of gain of information. En: Quarterly Journal of Experimental Psychology, vol. 4, pp. 11-36.

Hyman, R. (1953). Stimulus information as a determinant of reaction time. En: Journal of Experimental Psychology, vol. 45, pp. 188-196.

Hotho, A.; Jäschke, R.; Schmitz, C.; Stumme, G. (2006). Information Retrieval in Folksonomies: Search and Ranking. En: Lecture Notes in Computer Science, The Semantic Web: Research and Applications, Vol. 4011/2006, pp. 411-426.

Kules, W.; Wilson, M. L.; Schraefel, M.C.; Shneiderman, B. (2008). From Keyword Search to Exploration: How Result Visualization Aids Discovery on the Web. Technical Report 1516920080208, School of Electronics and Computer Science, University of Southampton. Disponible en: <http://eprints.ecs.soton.ac.uk/15169/>

Kuo, B.Y-L.; Hentrich, T.; Good, B.M., Wilkinson, M.D. (2007). Tag clouds for summarizing web search results. Proceedings of the 16th international conference on World Wide Web, Banff, Alberta, Canada, pp. 1203-1204.

Landauer, T.K.; Nachbar, D.W. (1985). Selection from alphabetic and numeric menu trees using a touch screen: Breadth, Depth and Width. CHI'85 Proceedings, ACM, April 1985, pp. 73-78.

Lehwark, P; Risi, S.; Ultsch, A. (2008). Visualization and Clustering of Tagged Music Data. En: Data Analysis, Machine Learning and Applications, Springer, pp. 673-680.

Klerkx, J.; Duval, E. (2009). Visualising Social Bookmarks. Journal of Digital Information, vol. 10, n. 2.

Mehlenbacher, B.; Duffy, T.M.; Palmer, J. (1989). Finding information on a Menu: Linking Menu Organization to the User's Goals. En: Human-Computer Interaction, vol. 4, pp. 231-251.

Nielsen, J. (2006). F-Shaped Pattern For Reading Web Content. UseIt Alertbox, 17 de Abril de 2006. Disponible en: http://www.useit.com/alertbox/reading_pattern.html

Provost, J. (2008). Improved Document Summarization and Tag Clouds via Singular Value Decomposition. Thesis submitted to the School of Computing in conformity with the requirements for the degree of Master of Science, Queen's University, Kingston, Ontario, Canada. Disponible en: <https://eggloop.its.queensu.ca/handle/1974/1476>

Risi, S.; Mörchen, F.; Ultsch, A.; Lewark, P (2007). Visual mining in music collections with Emergent SOM. Proceedings Workshop on Self-Organizing Maps (WSOM), Bielefeld, (2007). Disponible en: <http://www.mybytes.de/papers/risi07visual.pdf>

Rivadeneira, A. W.; Gruen, D.M.; Muller, M.J.; Millen, D.R. (2007). Getting our head in the clouds: toward evaluation studies of tagclouds. Proceedings of the SIGCHI conference on Human factors in computing systems, San Jose, California, USA, pp. 995-998.

Rosenfeld, L.; Morville, P. (2002). Information Architecture for the World Wide Web. 2nd edition. ISBN 0-596-00035-9. 2002.

Salton, G.; Wong, A.; Yang, C.S. (1975). A Vector Space Model for Automatic Indexing. Communications of the ACM, 18 (11), November 1975, pp. 613-620.

Schrammel, J.; Leitner, M.; Tscheligi, M. (2009). Semantically Structured Tag Clouds: An Empirical Evaluation of Clustered Presentation Approaches. Proceedings of the 27th international conference on Human factors in computing systems, Boston, MA, USA, pp. 2037-2040.

Schrammel, J.; Deutsch, St.; Tscheligi, M. (2009). The Visual Perception of Tag Clouds - Results from an Eyetracking Study. 12th IFIP TC13 Conference in Human-Computer Interaction INTERACT 09, Uppsala, Sweden, August 24-28, 2009.

Shergold, C.; Good, J.; Davies, J.; Gheorghiu, M. (2008). Using tag clouds to facilitate search: An evaluation. University of Sussex, Sussex, UK.

Shneiderman, B.; Plaisant, C. (2006). Strategies for evaluating information visualization tools: multi-dimensional in-depth long-term case studies. Proceedings of the 2006 AVI workshop on BEyond time and errors: novel evaluation methods for information visualization, Venice, Italy, pp. 1-7.

Sinclair, J.; Cardew-Hall, M. (2008). The folksonomy tag cloud: when is it useful?. Journal of Information Science, Vol. 34, n. 1, 2008, pp. 15-29.

Slingsby, A.; Dykes, J.; Wood, J.; Clarke, K. (2007). Interactive Tag Maps and Tag Clouds for the Multiscale Exploration of Large Spatio-temporal Datasets. 11th International Conference Information Visualization (IV'07), pp. 497-504.

Ware, C. (2003). Design as Applied Perception. En: Carroll, J.M. (Ed.). HCI Models, Theories and Frameworks: Toward a Multidisciplinary Science. Morgan Kaufman Publishers: San Francisco (USA). ISBN 1-55860-808-7.

Wolfe, J. M.; Cave, K. R.; Franzel, S. L. (1989). Guided Search: An alternative to the Feature Integration model for visual search. En: Journal of Experimental Psychology. Human Perception and Performance, vol. 15, n. 3, pp. 419-433.

Wolfe, J. M.; Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it?. En: Nature Reviews: Neuroscience, Vol. 5, n. 6, pp. 495-501.

Wolfe, J.M. (2006). Guided Search 4.0: Current Progress with a Model of Visual Search. En: MIT Scene Understanding Symposium, Spring 2006.

Recibido: 2009-09-09. Aceptado: 2010-06-11