

Representación temática y recuperación de los noticieros de televisión en la web: propuesta y aplicación de un modelo metodológico

Catalina Naumis Peña*
Silvano Soto Hernández**

Artículo recibido:
17 de febrero de 2018
Artículo aceptado:
18 de septiembre de 2018
Artículo de investigación

RESUMEN

Los noticieros se trasladan de la televisión a la web para acercarse a los usuarios. La representación textual de su contenido es importante para posicionar la noticia en la web en motores de búsqueda y facilitar su consulta y preservación. El valor económico que adquiere la difusión de los noticieros de la televisión incrementa cuanto más claro es el tema para las tecnologías web. En el presente estudio, basado en las metodologías del tratamiento documental y de la Investigación-Acción, se propone un modelo de representación temática de contenido para optimizar la recuperación a fin de promover el tráfico de usuarios. Se puso el modelo a prueba en un estudio de caso y se presentan los resultados obtenidos.

- * Instituto de Investigaciones Bibliotecológicas y de la Información, UNAM, México
naumis@unam.mx
** IDMS Consulting SC, México
silvano_soto@idms.com.mx

Palabras clave: Noticieros Televisivos en la Web; Representación Temática del Video Periodístico; Posicionamiento Web del Noticiero Televisivo

Subject area representation and retrieval of television news on the web: proposal and application of a methodological model

Catalina Naumis-Peña and Silvano Soto-Hernández

ABSTRACT

Newscasts migrate from television to the web in order to reach more users. The textual representation of news content serves to position the news on web in search engines and facilitate their consultation and preservation. The economic value of television news broadcasts increases with increasing clarity of the enunciation of the subject area for web technologies. Employing documentary management and action research methodologies, this paper proposes a subject area representation model to optimize both retrieval and user traffic. The model is put to the test in a case study and the results obtained are presented.

Keywords: Television News on the Web; Subject Area Representation of Journalistic Video; Web Positioning of Television Newscasts

INTRODUCCIÓN

La representación temática del contenido facilita la visibilidad, recuperación y acceso de los documentos a través de internet. En el caso de las noticias, la recuperabilidad se expresa a través de la consulta de los usuarios en la web o apps y, no menos importante, en la reutilización de los propios medios, sus editores y periodistas (Sobak y Pharo, 2017: 739).

Los teóricos que han estudiado la representación del contenido a través de la indización destacan cómo se orienta cada vez más hacia grupos de usuarios específicos (Hjørland, 2017: 59). En el ámbito de la edición y publicación en la web, la indización se ha especializado y enriquecido a la optimización para

buscadores o SEO (*Search Engine Optimization*). Al respecto, la investigación bibliotecológica especializada es escasa y, todavía más, tratándose de los contenidos audiovisuales (Arlitsch, O'Brien y Rossmann, 2013).

Curran (2004) señaló que el SEO influye positivamente en la publicación web a través de la manipulación de elementos utilizados por los buscadores para determinar su relevancia, y propuso la realización de una extensa investigación sobre palabras y términos clave para determinar cuáles son usados en títulos y descripciones. Luego, Curran (2004) y Raisinghani (2005) destacaron que existen oportunidades de mejora para aportar mayor puntaje a la evaluación de los sitios web; por ejemplo, el uso de palabras clave en el código. Tanto Curran como Raisinghani convergen con las recomendaciones de Zhang y Dimitroff (2005), quienes identificaron que al optimizar las etiquetas semánticas se alcanzan posiciones más relevantes en la búsqueda. Descubrieron que este fenómeno se repetía en sitios donde las páginas clave o descriptores temáticos aparecían en las etiquetas de título y en el cuerpo del texto.

Por otro lado, Sen (2005) desarrolló un modelo teórico para comparar los efectos del SEO con las estrategias de posicionamiento pagado *versus* el tráfico orgánico o natural de audiencia y Zing y Lin (2006: 519-529) reconocieron que, aunque el SEO estaba resultando de gran interés en la industria publicitaria, no ocurría lo mismo en el ámbito de la investigación científica y académica; menos en la industria televisiva. Más tarde, Malaga (2007: 68-82) estructuró una concepción sobre la planificación del SEO para incrementar la relevancia y la visibilidad.

El impacto del SEO ha sido tal que, en 2009, la *British Broadcasting Company* (BBC) desarrolló una posición laboral de especialista en SEO Periodismo y, en 2011, *Los Angeles Times* instituyó al SEO *Chief* (Ellis, 2011). Por su parte, para Stabiner (2011) y Bakker (2012), el SEO se puede comprender como un proceso que sirve a los editores web para crear y ajustar los encabezados o los textos de las noticias a fin de mejorar su posicionamiento.

Spyridou y Veglis (2012) aportan que las organizaciones periodísticas, entre otras actividades, redefinen y modifican sus procesos de producción alrededor del SEO para mejorar posicionamiento y tráfico, asunto que luego validaron Giomelakis y Veglis (2016) al concluir que existe una clara correlación entre resultados y optimización. Gonzalo, Codina y Rovira (2015) añaden que las tendencias se enfocan en dos aspectos relacionados con los usuarios: a) perfil del usuario y b) la intención temática de la búsqueda.

El campo de contribución de la bibliotecología y los estudios de la información en el enriquecimiento de la representación temática de los contenidos en la web tiene claras oportunidades en lo que se refiere a la investigación, la elaboración de diagnósticos y el diseño de intervenciones para la

mejora continua y la innovación. Texto, audio, video y animaciones conforman un discurso que, por su formato, demanda de un análisis documental especializado donde se consideren aspectos específicos como el visionado, el análisis del contenido, la elaboración del resumen, los descriptores temáticos con perspectiva semántica para el SEO, así como el etiquetado. En el caso de los noticieros televisivos en la web es necesario realizar un esfuerzo adicional, pues si bien existe un cuerpo de texto que acompaña al recurso multimedia, en realidad la mayor parte del contenido enriquecido se encuentra justo en dicho elemento. De ahí la utilidad de un modelo de representación y recuperación temática (Engerer, 2017: 661).

En Noruega, un estudio realizado en una televisora muestra las dificultades en la indización de contenidos televisivos de *Norwegian Broadcasting Corporation*. Los principales hallazgos revelan una indización incompleta, inconsistente y desigual (Sobak y Pharo, 2017: 739). En México, poco se conoce sobre estos y otros aspectos especializados en el discurso de los noticieros televisivos en la web. Por ello, se tomó como punto de partida el diagnóstico sobre el uso de métodos, técnicas o herramientas de representación temática en televisoras. Posteriormente se diseñó un modelo basado en experiencias de indización, con perspectiva bibliotecológica, que se probó en un entorno de negocio real para observar el efecto en el tráfico de usuarios.

METODOLOGÍA

Se acudió al método conocido como Investigación-Acción (*Action Research*) que Comfort (1985) y Kemmis, McTaggart y Nixon (2014) han estudiado, analizado y sistematizado. Su perspectiva fue tomada como punto de partida para construir la metodología, considerando las cinco etapas que guiaron el presente estudio. Los autores mencionados retoman el desarrollo atribuido a Lewin (1948) en la psicología social, seguido entre otros por Checkland (1981) y Checkland y Holwell (1998), cuyas principales contribuciones facilitaron la revisión del método con el rigor apropiado en la aplicación de la teoría de los sistemas de información y comunicación. Baskerville y Wood-Harper (1996: 235-246) describen la Investigación-Acción como un acercamiento científico basado en diagnósticos, intervenciones y generación de aprendizajes para construir un modelo cíclico que consta de cinco fases: diagnóstico, modelo metodológico, diseño de la intervención para probar el modelo, la implementación y evaluación correspondientes, así como la formulación de aprendizajes.

DIAGNÓSTICO

Para realizar el diagnóstico se establecieron criterios de análisis de la publicación de los noticieros en la web, considerando la interfaz del usuario y el código fuente de contenidos. El diagnóstico se realizó entre febrero y abril de 2015 en 14 televisoras mexicanas: Aprende Televisión Educativa, CNN en español, Canal Once, Efekto TV, Excélsior TV, Foro TV/Televisa, Fuerza Informativa Azteca, Milenio Noticias, MVS Noticias, Noticieros Televisa, Proyecto 40, Telefórmula, Televisión Metropolitana Canal 22 y TV UNAM.

Análisis de la interfaz de usuario. Se revisaron y analizaron las páginas web para identificar presencia o ausencia de metadatos de representación temática para el posicionamiento, la visibilidad y la recuperabilidad: título del noticiero, fecha y horario del noticiero, fecha y hora de posteo, conductores, síntesis del noticiero, palabras clave, nombres propios, temas relacionados, extensión de contenido en redes sociales, el noticiero en versión completa está disponible en vivo o bajo demanda, publicación de guion, escaleta, transcripciones, traducciones o subtítulos y publicación de fragmentos del noticiero bajo demanda.

Hallazgos del diagnóstico en la interfaz de usuario. Las televisoras cuentan con gestores de contenido preparados para asignar, de manera automática, título del noticiero, fecha del noticiero, horario de transmisión, fecha y hora de posteo y conductores. El 34 % no incluye una síntesis que exprese el tema, 86 % no contiene palabras clave o nombres propios como etiquetas complementarias, 79 % no agrega etiquetas sobre temas relacionados, 14 % no añade herramientas para compartir el contenido en redes sociales y 21 % no ofrece los noticieros en vivo o bajo demanda. En el 100 % está ausente la publicación de guiones o escaletas o transcripciones. Sólo un 7 % no ofrece fragmentos de los noticieros.

Análisis del código fuente. Por cada televisora o entidad productora se llevó a cabo la inspección del código fuente de las publicaciones de los noticieros disponibles, en vivo o bajo demanda, así como en cinco clips de video. Se tomaron en consideración los siguientes rasgos: densidad de palabras clave, etiqueta TITLE, etiquetas meta, etiquetas H1 y H2, análisis de palabras clave de la competencia, selección de palabras clave por estacionalidad, variantes idiomáticas y concatenaciones, etiquetado controlado y social, optimización de imágenes, metadescripciones, uso de microdatos y URL.

Hallazgos del diagnóstico del código fuente. Sólo 14 % presentó textos donde la densidad de las palabras clave superó el 15 %, colocándolas dentro del rango. Sin embargo, ninguna presentó una densidad mayor al 30 %. Es decir, los editores web publican “resúmenes periodísticos” pero carecen de

elementos semánticos suficientes para establecer el tema. En 71 % el código fuente no contiene palabras clave en la etiqueta <TITLE>, 57 % no suma enunciados que describan el tema y, en el 43 % restante, incorporan textos periodísticos con baja densidad de palabras clave (no obstante que sí lo hacen). El 100 % de las televisoras utilizan etiquetas <H1>, es decir, que colocan al menos un título en la página web, pero no siempre incluye alguna palabra clave temática. Por otra parte, ninguna utiliza las etiquetas <H2> para elaborar epígrafes, subtítulos o subsecciones en el cuerpo de los resúmenes periodísticos. Asimismo, no aprovechan la oportunidad de estudiar las palabras clave empleadas por sus competidores. En todos los casos se identificó la ausencia del uso de palabras clave seleccionadas por temporalidad; se agregan temas populares (*trending topics*) aunque este recurso no es capitalizado. También hay ausencia en el uso de variantes idiomáticas y concatenaciones.

Un 57 % de las televisoras ejecuta el proceso de asignar etiquetas; no obstante, se deduce la carencia de una estrategia clara para hacerlo, con base en la creación de listas, repertorios, vocabularios o taxonomías. Con un trabajo más estructurado destacan Efekto TV y Televisa que emplean palabras clave con un sentido más cercano a la indización temática. Sólo en el 21 % de los contenidos se encontró que las imágenes fijas o la primera imagen del reproductor de video no están optimizadas. En el 79 % de las televisoras no se identificó el uso de metadescripciones (uso de etiqueta <META = DESCRIPTION) para mejorar la descripción temática del contenido, a través de un resumen estructurado. Sólo aparece presente con un mayor grado de consistencia en Noticieros Televisa, Foro TV y Efekto TV. Este último caso sobresale porque utiliza una etiqueta específica: <META = ABSTRACT>. En el 100 % de las televisoras se identificó ausencia de microdatos mostrados en los *snippets* de las páginas de resultados arrojadas por motores de búsqueda. Por último, en el 71 % de las televisoras (10 casos) no se registran palabras clave en las URL.

PROPUESTA DE MODELO METODOLÓGICO

Una vez realizado el diagnóstico, en coincidencia con los trabajos de Serrano-Cobos (2015: 11), Malaga (2007: 72), Giomelakis y Veglis (2016, 22-33), así como Codina *et al.* (2016: 7), se presenta la propuesta de un modelo de representación temática de los noticieros en la web (*Figura 1*), tomando como base los factores esenciales para incrementar el valor de la optimización para buscadores (SEO).

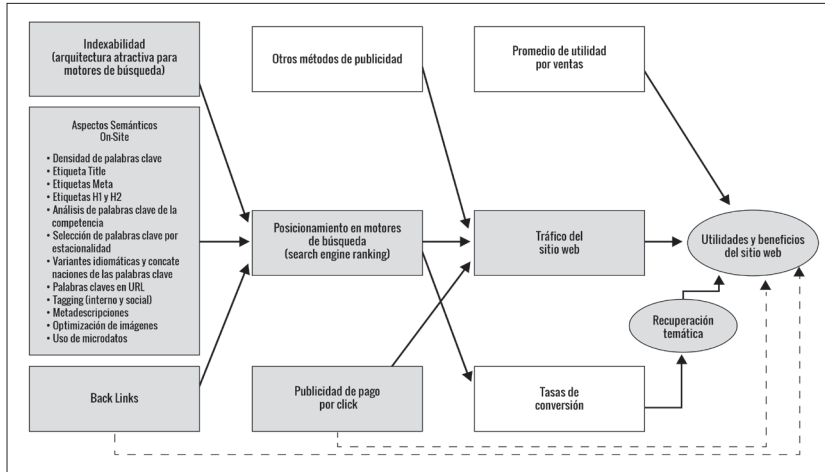


Figura 1. Modelo de representación temática del noticiero televisivo
 Con base en la aplicación de análisis y tratamiento documental del video, la recuperación de otros insumos de la creación del noticiero y aspectos semánticos on-site relevantes para SEO
 Fuente: Soto-Hernández (2017: 94) con base en Malaga (2007), Soto-Hernández (2009), Soto-Hernández y Naumis-Peña (2014) y Serrano-Cobos (2015)

El modelo es una aportación planteada desde la bibliotecología y los estudios de la información hacia el sector periodístico en la web. Tiene como objetivos los siguientes:

- Aplicar un modelo de análisis y tratamiento documental del video para refinar la representación temática (Soto-Hernández, 2009; Soto-Hernández y Naumis-Peña, 2014).
- Incrementar la relevancia de los contenidos digitales para los motores de búsqueda; su visibilidad, posicionamiento y acceso.
- Aprovechar el valor semántico en la web (Serrano-Cobos, 2015: 11).
- Facilitar la representación y recuperación temática mejorando la experiencia del usuario.
- Contribuir a la productividad.

En el modelo se incorporan los procedimientos técnicos del análisis y tratamiento documental del video: visionado, análisis de contenido, resumen, descripción temática y etiquetado. Cada uno de los aspectos semánticos más relevantes presentes en el modelo fueron explicitados en detalle y se presenta aquí la densidad de palabras clave como ejemplo.

- Representa el número de repeticiones en el cuerpo de un texto (no del total del sitio web).
- A través de la densidad de palabras clave, los motores de búsqueda determinan cuál es el tema principal o *aboutness* del contenido (Yablo, 2014: 1). Se recomienda utilizar todas las palabras de un mismo campo semántico incluyendo sinónimos, antónimos, variantes y formas verbales. De este modo aumenta la densidad, sin necesidad de incrementar las repeticiones.
- Es importante equilibrar la densidad de las palabras clave, así como su distribución en todos los espacios donde se puede aumentar la relevancia. De otro modo, el abuso puede ser interpretado por los motores de búsqueda como una mala práctica (*spam*), e incluso, ganarse una penalización en la evaluación de relevancia total del SEO.

De acuerdo con Soto-Hernández (2009: 101), el tratamiento documental del video engloba procedimientos y operaciones que son clave en la representación y recuperación temática de contenidos audiovisuales. Con este proceso técnico se busca satisfacer las demandas de la distribución, comercialización y visibilidad de los noticieros televisivos a través de un esquema multiplataforma (televisión, *web desktop*, *web mobile*, OTT como Netflix, app o medios sociales).

No obstante, las nuevas condiciones del contexto económico, social, empresarial, laboral y técnico, así como la velocidad de la obsolescencia tecnológica, han provocado la incorporación de profesiones emergentes. El etiquetado que pudieran realizar los lectores de las noticias tampoco es una solución. Por lo tanto, hay consenso en que las etiquetas generadas socialmente tienen problemas de calidad asociados con el uso de palabras no declaradas, polisemia, sinonimia y falta de jerarquía (Guy y Tonkin, 2006; Matusiak, 2006; Lu, Park y Hu, 2010), y la falta de distinción sobre a qué tipo corresponde una etiqueta (Springer *et al.*, 2008: 18). En el caso aún de la indización de imágenes fijas, los problemas existentes para indizar textos se multiplican (Matusiak, 2006: 294), debido a la riqueza semántica y la ambigüedad de la imagen (Melgar *et al.*, 2017: 350).

Cuando las empresas televisoras decidieron difundir contenidos en la web, surgió la necesidad de incorporar otro perfil profesional conocido como editor web cuyos objetivos y competencias son una obra en progreso. Éste es responsable de asegurar la ejecución de tareas que van desde la preparación de los materiales hasta otras vinculadas con el análisis de contenido.

El otro perfil profesional de reciente incorporación en las televisoras que publican noticieros en la web es el *SEO manager*, a cargo de supervisar

que cada contenido digital, como un sitio web, cuente con elementos que los motores de búsqueda exigen a través de sus algoritmos para alcanzar la mejor posición posible en las páginas de resultados.

Cabe preguntarse cuál es el lugar del bibliotecario, el profesional de la información, en este panorama (Arlitsch, O'Brien y Rossmann, 2013: 177-188), pues de las diferentes operaciones que deben realizar el editor web y el *SEO manager* hay algunas que, explícitamente, corresponden al ámbito de los procedimientos del campo bibliotecológico. La ejecución guarda apenas una distancia sutil pero relevante (Soto-Hernández, 2017: 101).

Para lograr que la representación temática se genere de forma efectiva, se proponen operaciones que proceden de la tradición bibliotecológica y de los estudios de la información, para los fines de recuperación, incremento de visibilidad, mejora en el posicionamiento en motores de búsqueda, experiencia de usuario y aumento en los indicadores de tráfico: resumen (no periodístico), descripción temática a través de palabras claves y etiquetado.

Las palabras clave son producto del análisis del contenido del video periodístico y extraídas de sus elementos textuales, expresados en el discurso del noticiero o del guion de una nota, reportaje, entrevista, etcétera. A diferencia de la indización especializada que se conoce en la bibliotecología, tratándose de documentos científicos, no resulta tan sencilla la extracción de palabras clave desde el título del noticiero (*v. g.* “CNN con Carmen Aristegui”) o de una noticia (*v. g.* “SSPCDMX aplicará operativo vial en Viaducto Tlalpan por obras”). Como lo señala Naumis-Peña (2007: 105), lo ideal es que sean representativas de los contenidos documentales y, en este sentido, cuando la extracción e indización se puede realizar por asignación y con participación humana, resultará más significativa.

En un artículo reciente se argumenta que las técnicas de inteligencia artificial no son lo suficientemente maduras para resolver el problema de la indización multimedia conceptualmente y no serán capaces de reemplazar a los indizadores humanos en el futuro previsible (MacFarlane, 2016: 180). El principal destinatario de las palabras clave es el motor de búsqueda y se persigue que haya la mayor congruencia entre el significado que están representando y la infinita gama de posibilidades cognitivas en una consulta efectuada por un usuario.

DISEÑO DE LA INTERVENCIÓN PARA APLICAR EL MODELO

Tomando como base los datos arrojados por el diagnóstico y el modelo de representación y recuperación propuesto en el apartado anterior, se estableció

contacto con las 14 televisoras analizadas. A todas se les propuso un programa de intervención. Finalmente, sólo una de las empresas aceptó que se ejecutara el proyecto; se reserva el nombre o marca de la organización por acuerdo de confidencialidad. Enseguida se presentan el caso y las acciones ejecutadas.

Reto

- Contrarrestar el efecto negativo en los principales indicadores de tráfico web que en febrero de 2015 disminuyeron 45 % (la aparición de Netflix y otro de tipo de programación fija) con respecto al año inmediato anterior, en la plataforma web y *web mobile* (web en dispositivos móviles).

Objetivos

- Diseñar una estrategia organizacional de tipo editorial y operativa para la representación temática.
- Implementar la estrategia de forma experimental en un ambiente de pruebas.
- Evaluar el impacto.
- Documentar aprendizajes.

La implementación. Etapas

- Establecimiento de un comité de trabajo.
- Diseño de la estrategia e indicadores.
- Implementación en un ambiente de pruebas.
- Análisis de datos.
- Formulación de aprendizajes.

Acciones

- Creación de comité de trabajo / Mayo, 2015.
- Diseño de la estrategia / Mayo, 2015.
- Análisis / Junio, 2015.
- Establecimiento de indicadores / Junio, 2015.
- Implementación en ambiente de pruebas / Julio-diciembre, 2015.
- Análisis de datos / Enero-febrero, 2016.
- Formulación de aprendizajes / Marzo, 2016.

Resultados esperados

- Frenar la caída del tráfico web y, de ser posible, incrementar los indicadores de desempeño.

- Indicadores observados: usuarios únicos absolutos, páginas vistas, videos vistos, tiempo de permanencia, perfiles demográficos.
- La ejecución del experimento se realizó en cinco noticieros que figuran dentro del portal web de la empresa. La intervención se aplicó en dos noticieros mientras que los otros tres sirvieron como medio de control, y se llevó a cabo de septiembre a noviembre de 2015.

EVALUACIÓN DE LOS RESULTADOS

Al concluir la intervención se procedió a la recolección, acopio y análisis de los datos relacionados con los indicadores clave de tráfico web establecidos. Los resultados se describen en la serie de figuras que aparecen a continuación. Como se aprecia en la *Figura 2*, al repunte de usuarios únicos absolutos que se presentó en febrero de 2015 siguió una caída constante que llegó a su nivel más bajo en julio, coincidiendo con el auge de Netflix y otros servicios que ofrecen programación fija. En febrero inició la etapa de implementación de la nueva estrategia de representación temática. Con las medidas, se logró un pico en la curva hacia octubre. Para noviembre, había comenzado a presentarse un comportamiento más estable con una ligera recuperación en diciembre, aun cuando la temporada navideña suele tener un impacto negativo en los indicadores de tráfico web. No se pagó campaña de impulso.



Figura 2. Usuarios únicos absolutos (millares)
Fuente: ComScore MyMetric, Google Analytics

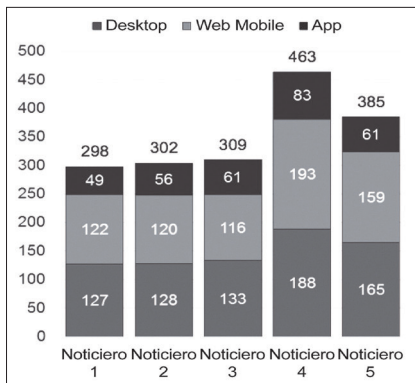


Figura 3. Usuarios únicos absolutos según noticiero (millares)
Fuente: ComScore MyMetrix, Google Analytics

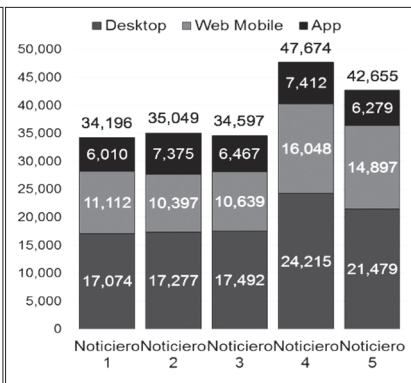


Figura 4. Páginas vistas según noticiero (millares)
Fuente: ComScore MyMetrix, Google Analytics

La Figura 3 expone el comportamiento que tuvieron los cinco noticieros que produce la empresa estudio de caso en el indicador de usuarios únicos absolutos. Sólo en los noticieros 4 y 5 (en los que se aplicó la política de representación) hubo un incremento notable. La diferencia parece sobresaliente en *web desktop* y *web mobile*; esto se debe a que la estrategia no se implementó en la app. En la Figura 4, en correspondencia, el número de páginas vistas también aumentó como resultado del experimento.

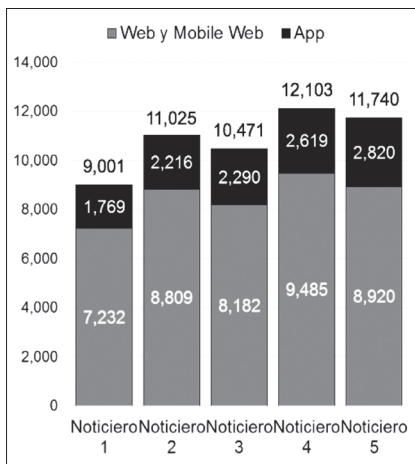


Figura 5. Videos vistos según noticiero (millares)
Fuente: ComScore MyMetrix, Google Analytics

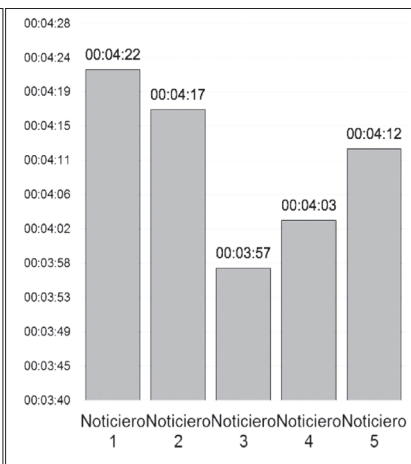


Figura 6. Tiempo de permanencia según noticiero (minutos/segundos)
Fuente: ComScore MyMetrix, Google Analytics

La *Figura 5* expresa la distribución del consumo de video según plataforma. Hubo crecimiento en los noticieros 4 y 5 por encima de los demás. En la *Figura 6* se observa el tiempo de permanencia realizando una actividad (*clicks*, recorrer la página, activar un reproductor de video); el incremento estuvo por arriba de la media comercial que ronda el 1.5 minutos en los noticieros donde se implementó la política de representación temática.

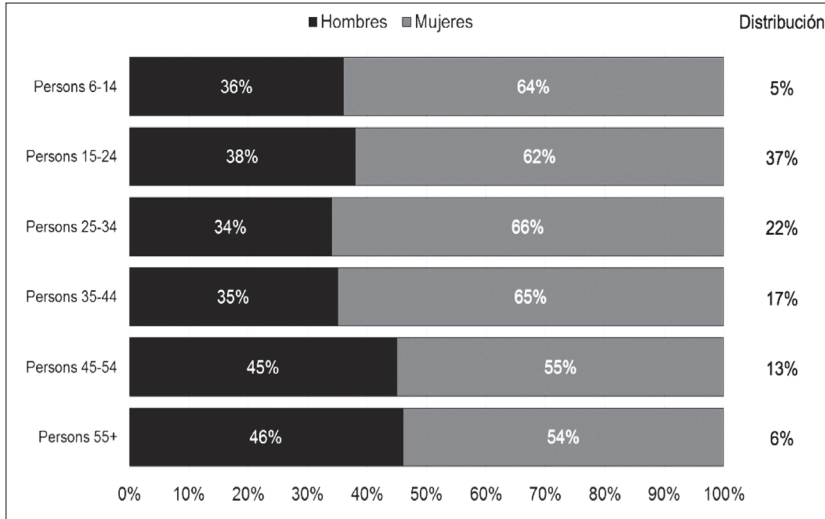


Figura 7. Perfiles demográficos generales (%)
Fuente: ComScore MyMetrix, Google Analytics

En la *Figura 7* se presenta el perfil demográfico de los usuarios. Se incorporó a niños y adultos mayores, lo que significa se alcanzó al llamado *target* comercial 4+ o familiar con el que, por cierto, antes no contaban. Esto aplica para los noticieros 4 y 5 en los que se realizó la intervención.

Formulación de aprendizajes

En la empresa estudio de caso se identificaron grandes dificultades para encontrar editores web con formación específica, profesionalizados y disponibles para llevar a cabo las tareas periodísticas y las de representación temática con rigor y efectividad. Ninguno de los editores web participantes en el proceso cuenta con algún tipo de entrenamiento especializado en biblioteología o gestión de información, o bien, en posicionamiento en motores de búsqueda desde una perspectiva semántica.

CONCLUSIONES Y RECOMENDACIONES

La televisión y la web están inmersas en la convergencia digital en el que se generan diversidad de expresiones creativas y donde se combinan estructuras, géneros, lenguajes y concepciones artísticas y periodísticas para satisfacer las necesidades de los usuarios, los anunciantes y el modelo de negocio.

Al probar el modelo propuesto para la representación de los noticieros en un entorno de negocio real, incluso en condiciones de estrechez presupuestal, se identificó el potencial que puede tener la contribución de la bibliotecología y los estudios de la información. Se pudo constatar que la metodología propuesta fue efectiva en un periodo breve de implementación.

Entre los hallazgos hay dos muy relevantes. El primero de ellos es que las prácticas de optimización para buscadores (SEO) se complementan y enriquecen con la representación temática resultado del tratamiento documental. El valor de la publicación digital en la que se difunden los noticieros televisivos incrementa cuanto más claro es el tema para los motores de búsqueda. Debería ser un objetivo el dotar de mayor recuperabilidad a la publicación web para desempeñar mejor sus funciones de comunicación social y, en su caso, los retos y metas comerciales de negocio que hacen posible la producción y distribución. Hay aquí una amalgama que se convierte en una vía de investigación, innovación y desarrollo para la disciplina bibliotecológica y los estudios de la información.

El segundo hallazgo está relacionado con las nuevas competencias profesionales que se requieren para gestionar el tratamiento documental como una innovación aplicada al sector industrial de la información en el entorno web. Hay demanda de recursos humanos preparados para realizar operaciones transaccionales, analíticas y estratégicas. En una era donde la visibilidad cobra un valor comercial, social, político, jurídico, histórico y cultural de enorme relevancia, hay un lugar importante para el bibliotecólogo como profesionista y para el investigador especializado.

Finalmente, el modelo muestra que el análisis documental y la representación temática y su aplicación al SEO del noticiero televisivo en la web requieren que el personal se actualice y profesionalice, traslapando su perfil con el del editor web en lo particular y con el de los profesionales de las industrias creativas en lo general (Mietzner y Kamprath, 2013: 280-294). Es decir, para el recurso humano implica diseñar y estructurar la publicación considerando la relevancia lingüística y semántica para mejorar la representación, recuperación, posicionamiento en los motores de búsqueda y experiencia de usuario.

En tanto que la información sea un elemento constitutivo del mercado y la recuperabilidad facilite el acceso a los usuarios o la recomendación para usarla, habrá necesidad de crear nuevas soluciones para posicionar los documentos y contenidos, con la finalidad de ponerlos en contacto con las personas.

REFERENCIAS

- Arlitsch, Kenning, Patrick OBrien y Brian Rossmann. 2013. "Managing Search Engine Optimization: An Introduction for Library Administrators". *Journal of Library Administration* 53 (2-3): 177-188.
- Bakker, P. 2012. "Aggregation, content farms and huffinization". *Journalism Practice* 6 (5-6): 627-637. doi:10.1080/17512786.2012.667266
- Baskerville, Richard. L. y A. Trevor Wood-Harper. 1996. "A Critical Perspective on Action Research as a Method for Information Systems Research". *Journal of Information Technology* 11: 235-246.
- Checkland, Peter y Sue Holwell. 1998. *Information, Systems and Information Systems Making Sense of the Field*. England: John Wiley and Sons.
- Checkland, Peter. 1981. *Systems Thinking, Systems Practice*, Chichester. England: John Wiley and Sons.
- Codina, Lluís, Mar Iglesias-García, Rafael Pedraza y Lucía García-Carretero. 2016. *Visibilidad y posicionamiento Web de informaciones periodísticas: El Framework SEO-RCP*. Barcelona: Serie Editorial DigiDoc-UPF. eRepositorio Universitat Pompeu Fabra. <https://www.upf.edu/web/digidoc/editorial-digidoc>
- Comfort, Louise K. 1985. "Action research: a model for organizational learning". *Journal of Policy Analysis and Management* 5 (1): 100-118.
- Curran, K. 2004. "Tips for Achieving High Positioning in the Results Pages of the Major Search Engines". *Information Technology Journal* 3 (2): 202-205. <https://doi.org/10.1108/10650750610706998>
- Ellis, Justin. 2011. "Traffic Report: Why pageviews and engagement are up at Latimes.com". *Nieman Journalism Lab*, August 16. <http://www.niemanlab.org/2011/08/traffic-report-why-pageviews-and-engagement-are-up-at-latimes-com/>
- Engerer, Volkmar. 2017. "Exploring Interdisciplinary Relationship Between Linguistics and Information Retrieval From the 1960s to Today". *Journal of the Association for Information Science and Technology* 68 (3): 660-680. <https://doi.org/10.1002/asi>
- Giomelakis, Dimitrios y Andreas Veglis. 2015. "Employing Search Engine Optimization Techniques in Online News". *Studies in Media and Communication* 3 (1): 22-33.
- Giomelakis, D. y Andreas Veglis. 2016. "Investigating Search Engine Optimization Factors in Media Websites". *Digital Journalism* 4 (3): 379-400. doi: 10.1080/21670811.2015.1046992
- Gonzalo, C., L. Codina y C. Rovira. 2015. "Recuperación de información centrada en el usuario SEO: categorización y determinación de las intenciones de búsqueda en la Web". *index.comunicación* 5 (3): 19-27.
- Guy, Marieke y EmmaTonkin. 2006. "Folksonomies: Tidying up tags?". *D-Lib Magazine* 12 (1). <https://doi.org/10.1045/january2006-guy>

- Hjørland, Birger. 2017. "Reviews of Concepts. Knowledge Organization". *Knowledge Organization* 44 (1): 55-64.
- Kemmis, Stephen, Robin McTaggart y Rondha Nixon. 2014. *The action research planner: doing critical participatory action research*. Singapore: Springer. <https://doi.org/0.1007/978-4560-67-2>
- Lewin, Kurt. 1948. *Resolving Social Conflicts*. New York: Harper.
- Lu, Caimei, Jung-Ran Park y Xiaohua Hu. 2010. "User tags versus expert-assigned subject terms: A comparison of LibraryThing tags and library of congress subject headings". *Journal of Information Science* 36 (6): 763-779. <https://doi.org/10.1177/0165551510386173>
- MacFarlane, Andrew. 2016. "Knowledge Organization and its Role in Multimedia Information Retrieval". *Knowledge Organization* 43 (3): 180-183.
- Malaga, Ross A. 2007. "The Value of Search Engine Optimization: An Action Research Project at a New E-Commerce Site". *Journal of Electronic Commerce in Organizations* 5 (3): 68-82. <https://doi.org/10.4018/jeco.2007070105>
- Matusiak, Krystyna K. 2006. "Towards user-centered indexing in digital image collections. OCLC". *Systems & Services* 22 (4): 283-298.
- Melgar Estrada, Liliana, Michiel Hildebrand, Victor de Boer y Jacco van Ossenbruggen. 2017. "Time-Based Tags for Fiction Movies: Comparing Experts to Novices Using a Video Labeling Game". *Journal of the Association for Information Science and Technology* 68 (2): 348-364.
- Mietzner, Dana y Martin Kamprath. 2013. "A Competence Portfolio for Professionals in the Creative Industries". *Creativity & Innovation Management* 22 (3): 280-294.
- Naumis-Peña, Catalina. 2007. *Los tesauros documentales y su aplicación en la información impresa, digital y multimedia*. México, Buenos Aires: UNAM, Centro Universitario de Investigaciones Bibliotecológicas, Library Outsourcing Service, Alfagrama.
- Raisinghani, Mahesh S. 2005. Future Trends in Search Engines". *Journal of Electronic Commerce in Organizations* 3: i-vii.
- Sen, R. 2005. "Optimal Search Engine Marketing Strategy". *International Journal of Electronic Commerce* 10 (1): 9-25.
- Serrano-Cobos, Jorge. 2015. *SEO. Introducción a la disciplina del posicionamiento en buscadores*. Barcelona: Editorial UOC.
- Sobak, Veslemoy y Nils Pharo. 2017 "Decentralized Subject Indexing of Television. Programs: The Effects of Using a Semicontrolled Indexing Language". *Journal of the Association for Information Science and Technology* 68 (3): 739-749. <https://doi.org/10.1002/asi>
- Soto-Hernández, Silvano y Catalina Naumis-Peña. 2014. "Análisis bibliotecológico de los noticieros televisivos mexicanos en la web". *El Profesional de la Información* 23 (1): 80-86. <https://doi.org/10.3145/epi.2014.ene.10>
- Soto-Hernández, Silvano. 2009. "Tratamiento documental del video: Propuesta metodológica". Tesis de maestría, Universidad Nacional Autónoma de México, Facultad de Filosofía y Letras.
- Soto-Hernández, Silvano. 2017. "Tratamiento documental: Representación y recuperación de los noticieros en la web". Tesis de doctorado, Universidad Nacional Autónoma de México, Facultad de Filosofía y Letras.
- Springer, Michelle, Beth Dulabahn, Phil Michel, Barbara Natanson, David Reser, David Woodward y Helena Zinkham. 2008. *For the common good: The Library of*

- Congress Flickr pilot project*. DC: Government of the United States; Library of Congress. http://loc.gov/rr/print/flickr_report_final.pdf
- Spyridou, P. y A. Veglis. 2012. "Sustainable Online News Projects: Redefining Production Norms and Practices". *Proceedings of the 10th World Media Economics and Management Conference*, Thessaloniki, May.
- Stabiner, K. 2011. "CJR Column Mentions The Simpsons. A second look at SEO". *Columbia Journalism Review* (March/April).
- Yablo, Stephen. 2014. *Aboutness*. Princeton, NJ: Oxford: Princeton University Press.
- Zhang, Jin y Alexandra Dimitroff. 2005. "The Impact of Metadata Implementation on Webpage Visibility in Search Engine Results", Part II. *Information Processing and Management* 41 (3): 691-715.
- Zing, Bo y Zhangxi Lin. 2006. "The Impact of Search Engine Optimization on Online Advertising Market". *Proceedings of the 8th International Conference on Electronic Commerce: The New W-Commerce: Innovations for Conquering Current Barriers, Obstacles and Limitations To Conducting Successful Business on the Internet*. New Brunswick, Canadá. <https://doi.org/10.1145/1151454.1151531>

Para citar este texto:

- Naumis Peña, Catalina y Silvano Soto Hernández. 2018. "Representación temática y recuperación de los noticieros de televisión en la web: propuesta y aplicación de un modelo metodológico". *Investigación Bibliotecológica: archivonomía, bibliotecología e información* 32 (77): 175-191. <http://dx.doi.org/10.22201/iibi.24488321xe.2018.77.57868>