
ISSN: 0719 – 0832



UTEM

Serie Bibliotecología y Gestión de Información N° 92, noviembre, 2014

Gestión digital de información de prensa

Christian Sifaqui



D · G · I

Departamento
de Gestión de
Información
Escuela de
Bibliotecología

Serie Bibliotecología y Gestión de Información es publicada desde Octubre de 2005 por el Departamento de Gestión de Información de la Universidad Tecnológica Metropolitana. Dr. Hernán Alessandri, 722, 6° piso, Providencia, Santiago, Chile, www.utem.cl

Sus artículos están disponibles en versión electrónica en Eprints in Library and Information Science: <http://eprints.rclis.org> y están indizados e integrados en la base de datos Academic Search Complete de EBSCO.

Está registrada en:

- Sistema Regional de Información en Línea para Revistas Científicas de América Latina, El Caribe, España y Portugal (LATINDEX)
- Ulrichs Web: Global Serials Directory.
- Dialnet, portal de difusión de la producción científica hispana,
- DOAJ – Directory of Open Access Journals
- Google Scholar.

Sitio Web: <http://bibliotecarios.cl/servicios/serie-bibliotecologia-y-gestion-de-informacion/>

Dirección Editorial

Carmen Pérez Ormeño, Directora Departamento de Gestión de la Información
Guillermo Toro Araneda, Director Escuela de Bibliotecología

Editor Jefe

Héctor Gómez Fuentes

Consejo Editorial

Académicos del Departamento de Gestión de Información

Mariela Ferrada Cubillos
Carlos Beltrán Ramírez

Presidenta del Colegio de Bibliotecarios de Chile A. G.

Gabriela Pradenas Bobadilla

Representante Legal

Luis Pinto Faverio

Decano Facultad de Administración y Economía

Enrique Maturana Lizardi

Autorizada su reproducción con mención de la fuente.

LAS IDEAS Y OPINIONES CONTENIDAS EN LOS TRABAJOS Y ARTÍCULOS SON DE RESPONSABILIDAD EXCLUSIVA DE LOS AUTORES Y NO EXPRESAN NECESARIAMENTE EL PUNTO DE VISTA DE LA UNIVERSIDAD TECNOLÓGICA METROPOLITANA.

Gestión digital de información de prensa

Christian Sifaqui¹

Ingeniero Civil en Informática.
Consultor independiente en temas de Informática Jurídica,
Procesamiento y Gestión digital de la Información.
sifaqui@gmail.com

Resumen

El artículo presenta una novedosa forma de gestión digital de la información proveniente de fuentes impresas (periódicos). Se presenta un breve desarrollo histórico del procesamiento de estas fuentes de información, describiendo el modelo lógico del desarrollo de un nuevo sistema de procesamiento de información, para finalmente presentar un prototipo que permitió resolver el relacionamiento de dos fuentes diversas en forma totalmente automática que permite el descubrimiento de nuevos patrones de información.

Abstract

The paper presents a novel way of managing digital information from print sources (newspapers). A brief historical development of the processing of these data sources is presented. Then the logical model of the development of a new information processing system is described. Finally presents a prototype that allowed resolving the relationship of two different sources in a fully automated way and also enabling the discovery of new patterns of information.

Palabras claves

Servicios de información, Procesamiento de periódicos, Digitalización, Indización, Archivo de prensa

Keywords

Information services, Processing of newspapers, Digitization, indexing, Press archive

¹ Trabajo realizado durante el periodo como Jefe de Sistemas y Servicios de Información en Red en la Biblioteca del Congreso Nacional de Chile.

Introducción

Uno de los pedidos más importantes y recurrentes de la comunidad del Congreso es contar con la mayor cantidad de información de su interés que haya sido publicada en medios de prensa nacional y regional, este tipo de requerimiento no sólo se circunscribe al ámbito parlamentario sino que ha sido de continuo interés para las bibliotecas por la información disponible en estos medios (Matoria, Upadhyay & Mishra, 2003; Gaur, Chand, Gaur & Singh, 2013).

Para resolver este requerimiento, la Biblioteca del Congreso Nacional (BCN) estableció en los años 50, dos unidades dedicadas a analizar y estructurar la información contenida en los diarios de circulación nacional, a saber:

- Unidad de Recortes de Prensa: analizaba las noticias de los diarios. Se recortaban físicamente con tijeras las noticias de interés, y se almacenaban estos recortes en sobres clasificados por materia, previa asignación de una tarjeta de identificación del recorte (ver figura 1). Actualmente son más de 4 millones de recortes de prensa los que existen en la bodega principal de la BCN (ver figura 2).
- Unidad de Indización de Prensa: esta unidad revisaba los diarios y redactaba fichas con un resumen de la noticia de interés, clasificando estas fichas bajo ciertos criterios y con una referencia para poder llegar a la noticia indizada (ver figura 3). Existen 15 ficheros con estas fichas (ver figura 4).

Por la información recopilada, estas dos unidades no tenían mayor interacción en la producción y en los resultados finales, pero los criterios para ambas unidades en su selección de las noticias, sí eran similares.

Estos criterios de selección, requerían recopilar noticias en las que fueran mencionados diputados, senadores, el Congreso, noticias del Poder Ejecutivo con trascendencia política o de agenda pública, noticias de temas culturales y temas internacionales que tuvieran relación con Chile o noticias de interés para las comisiones de RR.EE. del Congreso.



Fig.1. Recorte de prensa con su tarjeta de identificación



Fig.2. Sobres con recortes de prensa



Fig.3. Ficha de indización de prensa



Fig.4. Ficheros de indización de prensa

El trabajo de la Unidad de Indización de Prensa fue suspendido en 1990. La información producida por esta unidad (figuras 3 y 4) pasó paulatinamente al olvido.

Pero por el contrario, la Unidad de Recortes de Prensa continuó con su labor y a fines de los años 90 ésta fue potenciada con la puesta en marcha de un sistema computacional cliente-servidor (Invesdoc, ver figura 5), el que permitía automatizar el procesamiento de los recortes de prensa. Es interesante notar que proyectos similares se llevaron a cabo en los mismos años en otras partes del mundo (Mühlberger, 1999; Singh, Singh, Gupta & Kumar, 2000).

Esencialmente este sistema computacional emulaba el trabajo manual, ya que en vez de hacer el recorte en forma física, éste fue reemplazado por la digitalización de la noticia de interés. Personal de la unidad analizaba los diarios, marcaba las noticias con un lápiz y enviaba a un proveedor externo estos diarios marcados para la digitalización de las noticias seleccionadas y

posteriormente esta información era incorporada al sistema con metadatos específicos.



Fig.5. Sistema cliente-servidor Invesdoc

Noticias.bcn.cl

En el año 2008, la BCN decide reemplazar el sistema Invesdoc acorde a los objetivos estratégicos de:

- aumentar cobertura: ampliar el actual servicio a un mayor contenido por medios de prensa y aumentar la cantidad y el tipo de medios,
- autoservicio: contar con una interfaz de consulta web para traspasar el trabajo de búsqueda a los usuarios finales a través de una interfaz amigable y apropiada a sus necesidades.

Para lo anterior se diseñó y construyó un nuevo modelo de acopio y procesamiento de noticias de la prensa, delegando el acopio a un proveedor externo y programando algoritmos que permitieran acceder, buscar y recuperar estas noticias.

Lo anterior implicó un cambio conceptual mayor en cuanto a que el proceso de selección de los recortes iba a dejar de existir. Naturalmente lo anterior generó resistencia de las personas que usaban el sistema Invesdoc, debido que representaba un cambio de paradigma: Invesdoc funcionaba con un criterio de pre-coordinación, es decir, seres humanos decidían cuáles noticias iban a ser ingresadas y clasificadas en el sistema, mientras que Noticias (el nuevo sistema) funcionaría con un criterio de post-coordinación, lo que significaría que se ingresaría todo el diario y sería el usuario final el que - mediante una caja de búsqueda- definiría sus criterios de selección,

recibiendo como resultado las noticias que el sistema encontrara en base a estas consultas.

Noticias está operativo desde el año 2009, es de acceso web y se encuentra disponible sólo a la comunidad del Congreso² mediante la URL noticias.bcn.cl.

La concepción, desarrollo y puesta en marcha estuvo ejecutada en dos etapas principales:

- Primera etapa: diseño de un modelo de metadatos y contenidos de recortes de noticias con la flexibilidad de poder modelar noticias publicadas en la prensa escrita, así como en medios como internet, radio y televisión. En base a este modelo se realizó un proceso de adquisición de registros de noticias a un proveedor externo.
- La segunda etapa se dividió en tres procesos:
 - Modelamiento y programación de una aplicación que recibiera los datos desde el antiguo sistema Invesdoc. Esta fue una actividad realizada una única vez.
 - Modelamiento y programación de una aplicación que recibe los registros de prensa desde el proveedor externo, considerando el modelo de acopio y de almacenamiento indicado.
 - Modelamiento y programación de la indexación para consulta y provisión de servicios. Esta aplicación fue desarrollada usando Plone y el motor de búsqueda y relacionamiento Autonomy. Este módulo contempla una separación conceptual entre un módulo de Content Management y otro módulo de Content Delivery.

El módulo de Content Delivery de noticias.bcn.cl es capaz de agrupar sus resultados por medios y filtrar por fechas, incluso en su listado de aciertos (ver figura 6).

Una vez que el usuario accede a la noticia de interés, el sistema permite desplegar el contenido de los artículos (su texto, que proviene de un OCR) así como las imágenes de los recortes.

Este módulo permite, entre otras funcionalidades, hojear un periódico, extraer partes de una imagen, exportar noticias seleccionadas a un archivo pdf,

² Ley N° 17.336, Propiedad Intelectual, <http://www.leychile.cl/Navegar?idLey=17336>

simular un kiosco de diarios mostrando las portadas de los medios, visualizar noticias agrupadas en clúster por similitud de contenido.

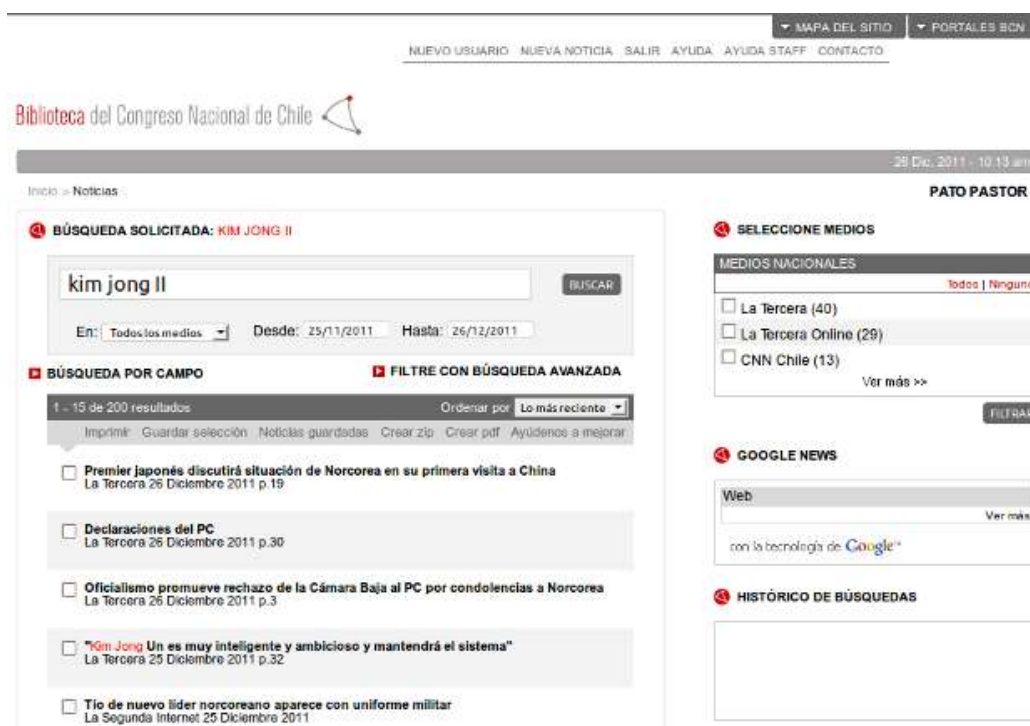


Fig.6. Listado de aciertos ante una consulta

Datos históricos

Uno de los objetivos secundarios, era poder gestionar en un solo sistema, las noticias que hubieran sido procesadas por la BCN, independiente de su formato físico. Por esta flexibilidad del modelo desarrollado para este nuevo sistema, es que se pudo dar inicio a un proceso de carga de los recortes históricos de la Unidad de Recortes de Prensa (1950 hasta 1999).

Para este proceso, se optó por realizar una selección del material acorde a criterios definidos por la Unidad Historia Política Legislativa.

El proceso se inicia seleccionando los sobres y los recortes contenidos en cada sobre, luego se preparan, se digitalizan e ingresan con sus metadatos al sistema, quedando inmediatamente disponibles para los usuarios, tal como lo muestra la figura 7.

Imprimir Enviar por correo Guardar en su cuenta Crear zip Crear pdf

EL VOTO FEMENINO ES DESDE AYER LEY DE LA REPUBLICA

La Nación 9 Enero 1949 p. 10-16

"La mujer humanizará la política chilena", dijo el Jefe de Estado

En brillante acto, el Presidente consagró derechos de la mujer
REFRENDARON MINISTROS DE INTERIOR Y JUSTICIA

En una nota de entusiasmo y emoción se tradujo el magnífico e histórico acto de la promulgación de la ley que otorga los derechos políticos a la mujer, efectuado a mediados de ayer en el Teatro Municipal de Santiago con asistencia de S. E. el Presidente de la República, Ministros de Estado, altas personalidades, parlamentarios y de una enorme concurrencia de mujeres que llenaba por completo las distintas aposentaduras de nuestro primer coliseo.

HERMOSO ESPECTACULO
Desde las 10 horas empezaron a llegar al Municipal mujeres de todas las clases sociales, que se fueron distribuyendo en palcos platea, anfiteatro y galería, ofreciendo un hermoso golpe de vista con sus trajes multicolores y con el entusiasmo que la gran victoria alcanzada ponía en sus rostros, y en las diversas manifestaciones de su espíritu que salpicaban el ambiente de fiesta que remaba en el Teatro.

LLEGA S. E. ACOMPASADO DE SU ESPOSA
A las 12:45 horas en punto llegaba al Teatro el Excmo. señor González Videla acompañado de su esposa, señora Rosa Markmann de González Videla, y del Edecán de servicio, Comandante Calderón. S. E. hizo su entrada por la puerta que da acceso directo al escenario, siendo recibido por el Ministro del Interior, don Immanuel Holger, a quien acompañaban sus colegas de Justicia, señor Luis Felipe Letelier; del Trabajo, don Ruperto Puga; de Vías y Obras, don Ernesto Merino Segura; de Educación, don Armando Mallet, y de la señora Ana Figueroa, presidenta nacional de la Federación de Instituciones Femeninas, y otras damas integrantes de la directiva de este organismo. En el escenario aguardaban al Primer Mandatario el Presidente del Senado, don Arturo Alessandri Palma; los senadores señores Pedro Opitz, Elodorio Domínguez y Salvador Allende; los diputados señores Miguel Luis Amunátegui y Carlos Mele; el Alcalde de Santiago, don José Santos Salas; las Sras. Amanda La Barca, Clara Williams de Finco, Georgina Purand, Aida Yávar de Figueroa, Raquel García de Zamorano, Berla Silva Hill de Müller y otras dignitas del movimiento femenino.

EL ACTO
Al levantarse el telón, el público prorrumió en entusiastas manifestaciones de simpatía para la persona del Excmo. señor González Videla y para su esposa, señora Rostita de González Videla, aulen ha sido la más decidida defensora de las aspiraciones políticas de la mujer chilena.

MILES DE VOCES ENTONAN EL HIMNO NACIONAL
El Teatro rindió en seguida un cariñoso homenaje al Excmo. señor González Videla. Puestos de pie la enorme concurrencia, miles de voces femeninas entonaron el Himno Nacional, que incluso fué cantado por S. E., su esposa y sus Ministros de Esta-Una salva de aplausos, que se prolongaron por espacio de varios minutos, puso término a la ejecución del Himno Patrio. A continuación el Orfeón de Carabineros ejeculó el Himno a la Mujer.

LOS DISCURSOS
En primer término habló la Presidenta Nacional de Instituciones Femeninas, señora Ana de Figueroa quien, visiblemente emocionada, pronunció un hermoso discurso, cuyo texto íntegro damos en información separada.

Recortes:




Fig. 7. Noticia/recorte de prensa histórica, mostrando OCR e imagen original.

Vinculando información

Como se indicara al inicio de este artículo, las fichas de la Unidad de Indización de Prensa habían quedado prácticamente quedado en el olvido para la BCN, excepto para las personas de la Sección de Prensa, hasta que visitas a las bodegas permitieron redescubrir estos ficheros, las fichas allí contenidas y la invaluable información allí incluida.

Unas de las muchas ventajas de estas fichas es que tienen un encabezamiento de materia, que hace muy expedito el acceso a la información allí contenida. Además -como muestra la figura 8- se tiene un rápido acceso a todas las noticias que tienen relación con alguna materia.



Fig. 8. Fichas con separadores de encabezamiento de materia

Por otro lado, al estar ya disponibles y procesados una gran cantidad de recortes de prensa históricos (de 1950 a 1999) en Noticias, es que se propuso desarrollar un prototipo para realizar una vinculación automática de las fichas con los recortes, tal como lo indica la figura 9, porque la conjetura era que se podría producir un nuevo espacio de interacción de la información disponible, al ofrecer un nuevo tipo de acceso.

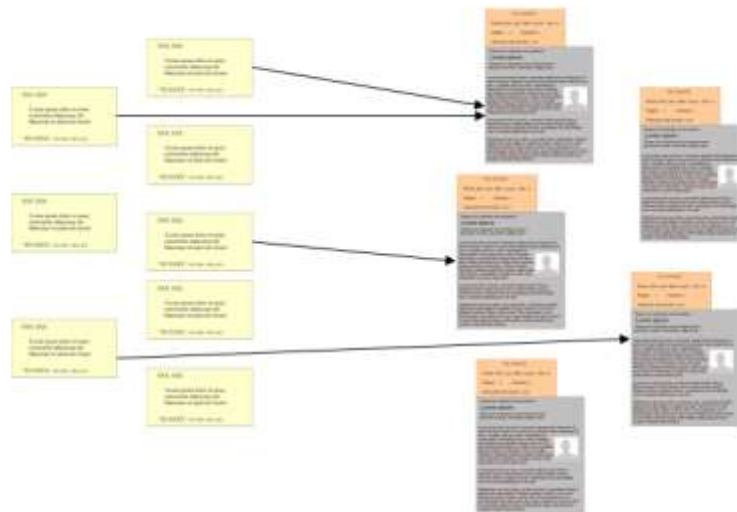


Fig. 9. Relación entre fichas de indización de prensa y recortes de noticias

Como se muestra en la figura 9, un recorte de noticia puede tener asociado varias fichas con diferente encabezamiento de materia, aunque también pueden existir recortes de noticias que tengan sólo una ficha asociada y también existirán fichas a las que no se les podrá asociar un recorte de prensa. Esto es así por al menos las siguientes razones:

- Discrepancia en proceso de selección de noticias por parte de la Unidad de Indización de Prensa y Unidad de Recortes de Prensa.

- Desaparición de fichas: se ha detectado que faltan fichas de indización.
- Recortes aún no digitalizados.

Prototipo

Las ideas para desarrollar un prototipo eran, por un parte confirmar la suposición anterior, es decir, poder relacionar digitalmente información que provenía desde distintas fuentes aunque relacionadas semánticamente y por otra poder conocer el esfuerzo y los tiempos involucrados en un espacio acotado de información y poder encontrar situaciones perfectibles con resultados concretos.

Los datos que se usaron para este prototipo fueron definidos por la información que digitalmente ya estaba disponible. Acorde a los criterios de la Unidad Historia Política Legislativa, ya se había iniciado un proceso de digitalización de recortes de prensa, que había comenzado con los sobres clasificados con la materia de “Partidos Políticos” y finalizados éstos, había proseguido con sobres con la materia “Partido Radical”.

Para poder construir el prototipo había que decidir por un período de tiempo y por una materia, para este efecto se eligió al Presidente del Partido Radical entre los años 1969 y 1973, Senador Anselmo Sule.

Se procedió a revisar los ficheros y recuperar las fichas de indización de prensa, proceso que obtuvo como resultado 107 fichas. El proceso de digitalización de esas 107 fichas tomó 10 minutos. El proceso de OCR posterior tomó 5 minutos sin corrección, pero dos personas consumieron 2:45 hrs. cada una para revisar los OCRs. Como aún permanecían errores, una persona revisó estas correcciones, proceso que demoró 1 hora app.

El prototipo fue implementado como una opción especial en el sistema Noticias y lucía como lo indica la figura 10.

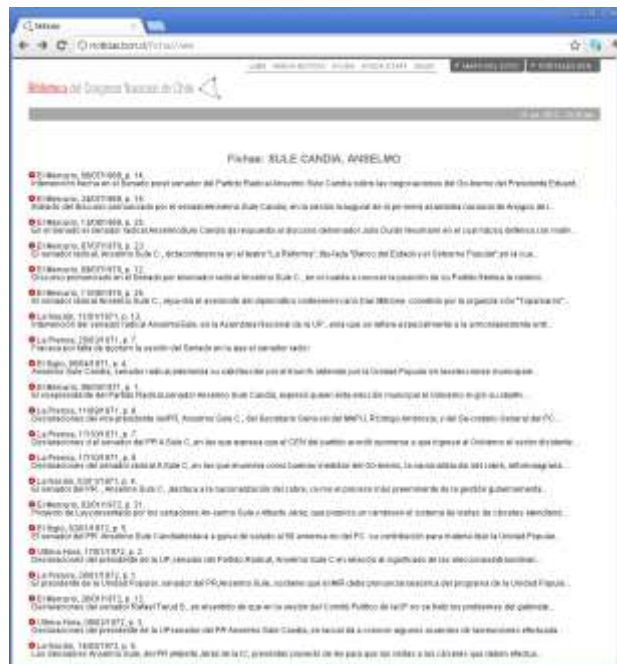


Fig. 10. Listado de las fichas

La figura 10 muestra el listado completo de las fichas digitalizadas ordenadas por fecha.

Al seleccionar una ficha (ver figura 11), se muestra el contenido de la ficha seleccionada y al mismo tiempo:

- se efectúa una búsqueda en los recortes de prensa, para encontrar el artículo al que hace referencia esta ficha, utilizando para esto, el medio, la fecha y la página. Puede ocurrir que no exista el recorte, que exista uno o que existan varios recortes que cumplan con esos criterios. En este último caso (en una página de periódico pueden haber varias noticias), para poder determinar con mayor precisión, se utiliza el texto de la ficha y se compara semánticamente con el texto del artículo para determinar si hacen referencia a los mismos conceptos,
- se buscan los recortes de noticias de ese día de otros medios cuyo contenido tenga relación semántica con el texto de la ficha,
- se buscan los recortes de las noticias de la semana que tengan relación con la ficha para poder entregar un contexto más amplio, ya que muchas veces las noticias respecto a un tema aparecen por varios días en los medios.



Fig. 11. Visualización de una ficha y recortes relacionados

El prototipo fue exhibido a funcionarios de la BCN, quienes ante la vista de las posibilidades que ofrecía, apoyaron el incrementar la cantidad de fichas a digitalizar, siendo seleccionadas las fichas de los siguientes congresistas:

- Gabriel González Videla: 460 fichas (1935 a 1981)
- Carlos Morales Abarzúa: 118 fichas (1951 a 1988)
- Raúl Rettig: 2 fichas (1949)
- Jonás Gómez: 33 fichas (1962-1969)
- Luis Bossay Leiva: 167 fichas (1949-1989)
- Enrique Silva Cimma: 110 fichas (1959-1990)

El prototipo quedó totalmente operativo en el sistema Noticias (ver figura 12) y se planifica para los años siguientes, el digitalizar más fichas de personalidades del partido Radical y Socialista, ya que actualmente se encuentra finalizado el ingreso de recortes del partido Radical y se está en pleno proceso de ingreso de recortes del partido Socialista.



Fig. 12. Prototipo estado a fines de 2012

Resultados inesperados

El proyecto pretendía vincular y relacionar dos fuentes de información exponiendo un nuevo tipo de acceso a las fuentes de información, pero surgieron dos resultados inesperados y relevantes: por un lado la ficha y la visualización de sus noticias relacionadas permiten detectar el hecho que es el fundamento de la noticia y por otro al tener a la vista los titulares y los textos se pueden ver explícitamente las diferencias en el tratamiento del hecho noticioso.

Identificación del hecho

Un resultado que no estaba contemplado se produjo visualizando la ficha que detalla una noticia del 13 de agosto de 1969 donde se da cuenta de un debate entre el Senador Sule y el Senador Durán (ver figura 11). Allí se puede ver cómo la prensa de ese día trató el tema y apreciar las diferentes ópticas de cada línea editorial y ver también las repercusiones que tuvo ese hecho tuvo durante la semana.

El episodio queda descrito acorde a la ficha redactada en la BCN de la siguiente manera: “En el Senado el senador radical Anselmo Sule Candia da respuesta al discurso del senador Julio Durán Neumann en el cual hizo su defensa con motivo de ser expulsado del Partido”.

La ficha hace referencia al artículo publicado en el Diario El Mercurio, el 13 de agosto de 1969 en la página 25. Este artículo lleva por título “Agrio Debate en el Senado entre Julio Durán y Anselmo Sule”.

Respecto de ese hecho, el mismo día 13 de agosto se publicaron los siguientes artículos de prensa:

- AL PR LE HAN ROBADO EL ALMA DE PARTIDO DEMOCRATICO, La Tercera, página 5
- Enérgica Respuesta de Senador Sule a Durán, La Nación, página 3
- Julio Durán trapeó el suelo con los dirigentes del PR, La Tarde, página 5
- Única alternativa del PR está junto a fuerzas antimperialistas, El Siglo, página 3
- NO HUBO BRILLO EN EL SENADO, Las Noticias de Última hora, página 3.

El día anterior, 12 de agosto de 1969, el diario la Nación publicaba una noticia que indicaba lo que iba a ocurrir en el Senado mediante el artículo “Senador Sule Responde hoy a Julio Durán”, página 5 y un día después (14 de agosto de 1969) el diario El Mercurio publicaba un artículo con el título “Violento Enfrentamiento De Ex Correligionarios”, página 25.

Por una parte es posible percatarse que de esta manera la ficha pasa a convertirse en el punto de inicio de algún hecho histórico y de esta manera poder analizar las derivaciones de ese suceso.

Lo anterior, superó la hipótesis que se había planteado originalmente, es decir, que al realizar la vinculación automática de las fichas con los recortes, podrían surgir nuevas experiencias de acceso, ya que lo que estaba ocurriendo era que se estaba obteniendo un hecho y se podría apreciar su evolución y tratamiento a través de las noticias de los diversos medios.

Detección de opiniones

Pero no sólo se obtuvo este resultado, sino que al visualizar los títulos de los artículos de prensa, se puede detectar una cierta opinión o sentimiento respecto del hecho acaecido, lo que queda reflejado al contrastar los títulos “Enérgica Respuesta de Senador Sule a Durán” versus “Julio Durán trapeó el suelo con los dirigentes del PR”.

Lo anterior lleva a pensar en la posibilidad de aplicar minería de opinión/análisis de sentimientos (Pang, Lee, Vaithyanathan, 2002; Pang & Lee, 2008) a estos artículos de prensa, para visualizar la polaridad de los sentimientos expresados. Pero como el texto de los artículos periodísticos está redactado en forma más neutral y las noticias de índole político poseen una expresión más sutil, es que podría ser más útil aplicar el framework propuesto por los autores (Khoo, Nourbakhsh & Na, 2012).

Conclusiones

Este prototipo permite ofrecer una nueva forma de acceder a información basándose en relacionar fuentes de información originalmente desunidas. Numerosas instituciones poseen acervos similares y el esquema que se describe en este artículo podría ser aplicable en escenarios similares.

Los motivos originales y en particular los resultados inesperados descubiertos, instaron a seguir digitalizando las fichas y los recortes para aprovechar todo el potencial de la información que está contenida en esos documentos.

Se estima que aplicando algoritmos de minería de opiniones para determinar la actitud de las noticias ante ciertos hechos podría ser una herramienta interesante para investigadores en estos temas.

Se espera que la descripción del desarrollo histórico del tratamiento de este tipo de fuente de información y de cómo en la BCN se ha ido avanzando en los métodos de procesamiento, así como la descripción de un prototipo, podría ser útil para el diseño e implementación de servicios de noticias para otras bibliotecas y así hacer disponible este acervo tan excepcional y útil.

Bibliografía

1. Gaur, Ramesh C.; Chang, Mahesh; Gaur, Kavita; Singh Yadav, Amar. Online newspaper clippings & news services for libraries: Experiences in Indian Libraries. En: IFLA WLIC 2013, Singapur, 2013.
2. Khoo, C.S.G.; Nourbakhsh, A.; Jin-Cheon Na. Sentiment analysis of online news text: a case study of appraisal theory. **Online Information Review**: 36(6): 858-878, 2012.
3. Muhlberger, Gunter. Digitisation of Newspaper Clippings: The LAURIN Project. **RLGDigi News**, 3(6): 1-20, 1999.
4. Pang, Bo; Lee, Lillian; Vaithyanathan, Shivakumar. Thumbs up?: sentiment classification using machine learning techniques. En: ACL-02 conference on Empirical methods in natural language processing, Praga, 2002. Proceedings of the 7th Conference on Empirical Methods in Natural Language Processing, 2002, p. 79-86.
5. Pang, Bo; Lee, Lillian. Opinion mining and sentiment analysis. **Foundations and Trends in Information Retrieval**: 2(1/2): 1-135, 2008.
6. Ram Kumar Matoria; P.K. Upadhyay; Arvind Mishra. NewsNIC: a Web-based, full-text news clipping service from the National Informatics Centre Library in India. **Program**, 37(3): 181-189, 2003.
7. Singh, Manoj; Singh Sanjay K.; Gupta, Rajiv; Kumar, Vijai. E-clippings: paperless nuclear science and technology news retrieval and archival application. En: Library and information networking, NAACLIN 2000: national convention on library and information networking, Madras, Chennai, 22-25 diciembre, 2000. Proceedings of the DELNEJ Conference, 2000. p. 129-137.

Serie Bibliotecología y Gestión de Información.

Títulos publicados 2014

N° 86 Tecnología para la transparencia legislativa. Romina Garrido Iglesias.

N° 87 Vinculación con el Medio y Biblioteca Universitaria. Yasna CatalánChávez

N° 88 Prácticas Lectoras de Jóvenes Infractores de Ley: Estudio de caso Centro Semicerrado La Cisterna. SENAME. Carolina Álvarez, Amanda Guajardo, Nicolás Álvarez

N° 89 Panorama de las bibliotecas municipales francesas: Visión de un ciudadano. Al Narea.

N° 90 Archivística para Bibliotecarios: Puente conceptual entre dos mundos. Wenke Adam

N° 91 Ontología de parlamentarios chilenos autores: Modelamiento y aplicación. Ángela Karina Quiroz Ubierna

N° 92 Gestión digital de la información de prensa. Christian Sifaqui

Disponible en :

<http://eprints.rclis.org>

NORMAS DE PUBLICACION

• **Objetivos**

La **Serie Bibliotecología y Gestión de Información** tiene por objetivo difundir la productividad, académica, las investigaciones y las experiencias de profesionales del área de la de Bibliotecología y Ciencia de la Información y del sector afín al mundo del libro y la lectura.

• **Alcance y política editorial**

Los trabajos a ser considerados en la Serie Bibliotecología y Gestión de Información, deben ser inéditos, no publicados en otras revistas o libros. Excepcionalmente el Comité Editorial podrá aceptar artículos que no cumplan con este requisito.

- **Arbitraje:** Los artículos recibidos serán sometidos a evaluación, a recomendación del Director de la Serie, donde el Comité Editorial enviará los trabajos a árbitros independientes para su aceptación o rechazo. En este último caso, se emitirá un informe al autor/a donde se señalen las razones de la decisión. El Comité Editorial podrá solicitar trabajos a autores de reconocido prestigio, quienes no serán sometidos al proceso de evaluación por árbitros.

• **Forma y preparación de artículos**

- **Extensión:** El artículo deberá tener una extensión entre 12 y 100 páginas, tamaño carta, espacio 1,5, cuerpo 12, incluidos gráficos, cuadros, diagramas, notas y referencias bibliográficas.

- **Idiomas:** Se aceptan trabajos en castellano, portugués e inglés, los cuales serán publicados en su idioma original.

- **Resumen y palabras claves:** El trabajo deberá tener un resumen en español e inglés en la primera página, de no más de 200 palabras, que sintetice sus propósitos y conclusiones más relevantes. De igual modo, deben incluirse tres palabras claves, que en lo posible no se encuentren en el título del trabajo, para efectos de indización bibliográfica.

- **Nota biográfica:** En la primera página, en nota al pie de página, deben consignarse una breve reseña curricular de los/as autores/as, considerando nacionalidad, título y/o grados académicos, desempeño y/o afiliación profesional actual y sus direcciones de correo electrónico, para posibles comunicaciones de los/las lectores/as con los autores/as.

- **Referencia bibliográfica:** Utilizar para las referencias bibliográficas la modalidad de (Autor, año) en el texto, evitando su utilización a pie de página. Ejemplo: (González, 2006). Agregar al final del texto, la bibliografía completa. Sólo con los/las autores/as y obras citadas, numeradas y ordenadas alfabéticamente. Para el formato de la bibliografía, utilizar la "Guía para la presentación de referencias bibliográficas de publicaciones impresas y electrónicas" disponible en formato electrónico en :

<http://eprints.rclis.org/6944/>

- **Derechos:** Los derechos sobre los trabajos publicados, serán cedidos por los/as autores/as a la **Universidad Tecnológica Metropolitana**.

- **Investigadores jóvenes:** El Comité Editorial considerará positivamente el envío de trabajos por parte de profesionales y/o investigadores/as jóvenes, como una forma de incentivo y apoyo a quienes comienzan su carrera en investigación.

- **Envío de Documentos**

Todas las colaboraciones deberán ser enviadas al correo electrónico: hector.gomez@utem.cl , en programa Word (office).