

El periodista y la documentación digital: nuevo espacio para la investigación

Mabel Calderín
Miladys Rojano

Es el momento, si no de descubrir, sí de enfatizar el peso específico que asume la documentación en el periodismo digital, puesto que se trata de un auténtico valor añadido que lo posiciona favorablemente con respecto a los medios tradicionales. (...) la brecha social se agudiza por la fractura informativa que se abre entre quienes tienen acceso a la información y los que no, y todavía más entre quienes son capaces de gestionar el caos informativo y los que son arrastrados por el diluvio de datos de la red

María Rubio Lacoba

El concepto de documentación digital es una referencia natural hacia Internet. Aunque ésta no es la única fuente que atesora información en formato digital, sí se considera el espacio de consulta y acceso a la información más rápido y directo del mundo contemporáneo. Sus límites, así como su tamaño son difíciles de calcular dado su crecimiento exponencial. Sin embargo, un estudio realizado por Gulli y Signorini (2005) sugiere que hay más de 11,5 billones de páginas solo en la *internet indexable*, esta es la parte de la web que es recuperable a partir de los motores de búsqueda generalistas como son Google, Altavista o Yahoo, entre otros.

Así, hay que considerar la existencia de dos niveles en Internet, uno visible que se conoce como la "Internet superficial" y uno oculto llamado "Internet invisible" o "web profunda". El primer nivel es accesible a través de los motores de búsqueda tradicionales y normalmente reúne contenidos en páginas estáticas, mientras que la localización de documentos en la web profunda está condicionada por la interrogación dentro de los propios sistemas manejadores de Bases de Datos (BD) particulares. Sin embargo, la frontera entre estos dos niveles no está claramente establecida porque en la web superficial se pueden encontrar documentos de la web profunda y viceversa (Salazar, 2005).

Una descripción aproximada del tamaño de este docuverso¹ se presenta en un estudio realizado por la Bright Planet Corporation en el año 2000, donde se confirma que la información pública en la web profunda es aproximadamente 500 veces mayor que en la web superficial. Y el tamaño total de esta web oculta alcanza los 550 billones de documentos individuales, mientras que la web visible contiene apenas un (1) billón de documentos (Bergman, 2001).

Por otra parte, la investigación como tarea sustantiva para la profesión periodística, es una de las labores que más cambios registra a partir de la incorporación de la tecnología digital en la salas de redacciones. El salto hacia los procesos informáticos abrió nuevas posibilidades de documentación, y con ello otras prácticas que contemplan desde estrategias de búsquedas distintas hasta un

¹ Según Theodor Nelson (1970) el docuverso es un mar de documentos relacionados todos mediante enlaces hipertextuales y completamente disponibles.

profesional con mayores capacidades para discernir sobre lo verdaderamente útil en un medio caracterizado por la abundancia de información.

La profesión periodística se encuentra ante nuevos desafíos relacionados con la capacidad para formular preguntas que puedan contribuir a descifrar la complejidad de los problemas contemporáneos; desarrollar y apropiarse de las herramientas para la búsqueda y validación de la información en las nuevas ofertas derivadas de la red de redes y desplegar el mayor esfuerzo analítico de la ampliación de los volúmenes de información disponibles y por último, convertir todos estos esfuerzos en un relato que se ajuste a las características de los medios y despierte el interés de los lectores y audiencias.

Al mismo tiempo, el escenario actual está basado en dos tendencias: la propensión a la digitalización de todo lo “digitalizable” como imágenes, sonidos, textos; y la necesidad de colocarlo en Internet (Codina, 2000).

Este panorama invita al profesional de la información a explorar nuevos soportes, específicamente la World Wide Web (WWW), para consultar BD de patrimonios culturales de distintas áreas; además de plantear el reto de resolver conflictos como exceso de contenido poco útil y la confiabilidad de las fuentes.

De acuerdo con Cervera (2006), “criterio y atención” son los elementos principales para investigar en la Red. Criterio para seleccionar los datos valiosos y desechar las manipulaciones. Atención para atraer a receptores dispersos en millones de bits.

La investigación en línea es una actividad relativamente nueva que involucra tanto a documentalistas como a periodistas, profesionales que deben conjugar áreas de conocimiento para extraer mejores beneficios de la Red.

Con la reconfiguración de perfiles y áreas de trabajo, el comunicador actual debe incorporar en sus labores algunas técnicas del proceso documental para optimizar actividades como recuperación de la información y planificación estratégica de la arquitectura de la información².

En ese sentido, Codina (2000) acota que ahora documentalistas y periodistas comparten nuevas tareas que pueden clasificarse en cuatro aspectos:

- Búsqueda y obtención de la información en la Red, distribuidores y BD especializadas (la “Internet invisible”).
- Evaluación, selección y descripción de recursos digitales.
- Procedimientos para mejorar la visibilidad en la web.
- Procedimientos característicos de la producción y distribución de informaciones.

Antecedentes de la documentación digital en el periodismo

Antes de *Google* y del surgimiento de otros motores de búsquedas actuales que se sumergen en la información superficial existente en la web, la investigación se desarrollaba en archivos físicos que se fueron transformando en BD digitales

² Según el glosario del Argus Center for Information Architecture, la Arquitectura de la Información (AI) es el “arte y la ciencia de la organización de la información que ayuda a la gente a satisfacer efectivamente sus necesidades de información (...) envuelve investigación, análisis, diseño e implementación” (González, 2003. p.2)

con acceso limitado, hasta llegar a convertirse en recursos informativos disponibles en Internet.

La evolución tecnológica permitió fundamentalmente mejorar la organización de materiales analógicos, y el desarrollo de potentes herramientas para buscar y recuperar información.

Los primeros mecanismos datan de los años cincuenta, cuando el acceso a un documento dependía de establecer conexión con un computador remoto que tuviera alojada la BD, a través del conocimiento del IP específico.

Salazar (2005) explica que esta situación mejoró cuando crearon un servicio de archivos centralizados o FTP³ anónimos que permitían el intercambio de documentos en directorios específicos.

Para esta época aún era necesario conocer el contenido alojado en el servidor específico. A partir de la apertura de la Red hacia otros sectores, distinto al académico y militar, se crea en 1990 *Archie*, la primera herramienta de búsqueda incluida en servidores FTP que recuperaba información y la depositaba en una BD central que se actualizaba una vez al mes.

Posteriormente, se creó *Gopher*, la cual generó el avance de permitir búsquedas por categorías y acceso a texto completo. Su aparición se considera el primer antecedente de los buscadores actuales como *Yahoo*, ya que “ofrecía al usuario una visión general y estructurada de toda la información de la que disponía; además, daba la oportunidad de pulsar en enlaces de hipertexto para moverse a áreas más especializadas” (Salazar, 2005, p.40).

A la par del desarrollo de estas herramientas y del concepto de Internet, se comenzó a generar documentación digital para resguardar legados históricos y permitir su acceso en línea. La primera en sumarse no fue la empresa periodística, los antecedentes iniciales se encuentran en recursos electrónicos académicos que proveían material a universidades y posteriormente a periodistas interesados en enriquecer sus indagaciones.

Entre las más importantes en el área de las ciencias sociales se encuentran *Academis Press*, *Blackwell Publishing*, *Brill Academia Publisher*, *Cambridge University Press*, *Ebsco Information Server*, entre otras. Su misión consiste en ser vehículos de difusión de la investigación científica de utilidad.

Concretamente en el campo periodístico, los medios impresos apenas utilizaban las BD iniciales, sobre todo porque eran muy especializadas, o por el costo que representaba para las empresas editoriales la consulta en algunas privadas. Existían entonces, y hoy en día más, algunas BD gratuitas, pero la diferencia informativa radica en disponer de esa información que tienen todos y de otras que son muy difíciles de conseguir.

Estados Unidos es uno de los países donde el valor de las BD viene determinado en función de su gestión informativa. Uno de los primeros medios impresos que apostó por la creación de BD fue *The New York Times*, el cual creó el *Information Bank* con millones de artículos de ese diario y de otros norteamericanos.

³ File Transfer Protocol (FTP) es uno de los protocolos de Internet que permite subir y descargar documentos de un servidor.

En Europa, los periódicos tardaron mucho tiempo en confeccionar BD. Inicialmente fueron las agencias de prensa las que decidieron desarrollarlas a partir de sus teletipos o cables diarios, para luego ofrecerlas como recurso de información a los diferentes medios. Entre las primeras agencias que se sumaron a esta iniciativa se encuentran France Press (Francia), ANSA (Italia) y EFE (España).

A la iniciativa de las agencias se sumaron las empresas que organizaron BD a partir de análisis de los artículos publicados en prensa. Una de ellas es la empresa española de servicios de teledocumentación Baratz. Los periódicos, por su parte, se interesaron luego de que evolucionara el concepto de Internet, si bien todos atesoraban en archivos privados importantes recursos informativos, su colocación en la Red y posterior venta, se debió al descubrimiento de un nuevo nicho comercial que hacía posible el acceso a valiosos documentos.

Bajo la nomenclatura de "Archivos", "Hemerotecas", "Siete fechas", "Siete días", "Centros de documentación", las versiones digitales de los diarios adoptaron distintos modelos para ofrecer sus colecciones.

En Estados Unidos el proceso comenzó en los años noventa con diarios como *USA Today*, cuyo modelo es seguido por la mayoría de los periódicos norteamericanos que incorporaron casi desde su lanzamiento motores de búsqueda hacia sus hemerotecas privadas.

En Europa, *Le Monde* de Francia gestionó sus archivos permitiéndoles a los usuarios acceder a la información gratis por un lapso de siete días y luego cobrar por ella. Su sistema actual reúne artículos desde 1987 y cuenta con más de 865.723 textos.

Los medios digitales en castellano incorporaron archivos en sus ediciones un tiempo después que los periódicos ingleses y franceses. En España, el diario *El País* permitió desde su lanzamiento en 1996 un acceso a sus ediciones de los últimos siete días, luego ese concepto varió en el año 2002 cuando incorporaron una hemeroteca con artículos que datan de 1976. En la región de Cataluña, en España, *La Vanguardia Digital* incluyó la sección 'Hemeroteca centenaria' en el año 2000. Y el diario *El Mundo* maneja dos tipos de archivos, el digital desde el año 2002 y el impreso con artículos desde 1994.

En Venezuela, la tendencia inicial se basó en consumir las BD y archivos digitales organizados en el exterior. Son muchos los diarios que encontraron en ellos aliados para la investigación de ciertos temas, sobre todo de carácter internacional.

Para el año 2000, las versiones de los diarios *El Nacional* y *El Universal* ya contaban con hemerotecas digitales. No obstante, la estrategia comercial se abordó de modo distinto. *El Nacional* comenzó con un modelo gratuito, pero ante el número de solicitantes del servicio lo ofrece en la actualidad por suscripción. *El Universal*, por su parte, permite la consulta gratuita de sus archivos desde 1996, fecha en la que incursionó en la Red.

Fuentes electrónicas como recursos de investigación para la comunicación

Como recomiendan los teóricos de la comunicación, la redacción periodística siempre debe ser concisa, clara, correcta, con rigor informativo,

interesante, rápida y universal (Martínez Alberto y Santamaría, 1993; Garrido 2006). No obstante, la transparencia en esta labor requiere algo más que consultar información y acercarle esos datos al público. Con una audiencia formada en las nuevas herramientas y con más posibilidades de participar en procesos informativos, el periodista actual debe garantizar una correcta interpretación de los datos y la eventualidad de que su noticia sea auditable.

El hecho de que el lector pueda acceder a las mismas BD que alimentaron la redacción de una nota informativa, y comunicarse con las fuentes exigirá entonces prácticas periodísticas más cuidadosas en el entorno digital. No basta con navegar en la Red, sino sumergirse para extraer contenido confiable.

Una aguja en el pajar sería prácticamente imposible de encontrar sin el uso de técnicas adecuadas: un gran imán, por ejemplo. Esto mismo ocurre en Internet, donde cada brizna de información de ese inmenso 'pajar virtual' está disponible para cualquier usuario que sepa buscarla. Tan sólo hay que saber la técnica adecuada en cada caso, y no conformarse con usar únicamente las herramientas y métodos más conocidos (Salazar, 2006, p.31).

Estrategias de búsquedas y recuperación de la información en línea

La búsqueda de respuestas verdaderamente relevantes a las necesidades de información de los comunicadores es un problema que se relaciona directamente con el tamaño cada vez mayor de la WWW. De acuerdo con Lyman y Varian (2000 cp Tramullas, 2002) cada año se produce en el mundo entre 1 y 2 Exabytes (1 Exabyte = 1 billón de Gigabytes, 250 Megabytes por persona) de información en todos los soportes. De esa información la mayor parte está registrada en soportes magnéticos y sólo el 0,003 % del total está impresa.

Hasta hace muy poco tiempo la difusión del conocimiento se limitaba al formato impreso, los manuscritos o la transmisión oral, pero con Internet surgió el nuevo formato de comunicación y acceso al conocimiento humano de la "sociedad de la información". Esta sociedad, marcada por el consumo creciente y acelerado de la información, impone el desarrollo de habilidades y destrezas en materia de búsqueda y recuperación de información realmente necesarias, especialmente para los profesionales de la información.

Rubio (2000) explica las tareas inherentes a los profesionales de la información a partir de cuatro funciones fundamentales:

1. Función selectiva-valorativa, basada en recurrir a fuentes web confiables por criterios de legitimación como su autoría, trayectoria, análisis de títulos, direcciones.
2. Función de recuperación, que incluye el modo de obtener información, estrategias de búsquedas y la verificación o comprobación de los contenidos.
3. Función gestora relacionada con la organización y diseño de la información para entornos digitales.

4. Función difusora caracterizada por la elaboración de contenido complementario que sirva de apoyo a la información noticiosa de última hora.

El desarrollo de cada una de ellas derivará en una mayor precisión en la actividad periodística, además de evitar la valoración simplista de que "todo lo que se consigue en la Web es basura". Si bien, no se trata de invadir áreas de trabajo, la tendencia actual se basa en una sinergia entre documentación y periodismo que comenzó a gestarse un tiempo atrás.

La búsqueda de información en la web no es una actividad exacta que se rige por un modelo único, sino que existen distintas maneras de llegar a obtener la respuesta deseada. No obstante, hay una serie de reglas que cuando se aplican de manera correcta incrementan el nivel de eficiencia y efectividad en la satisfacción de la necesidad de información.

Cuando se busca información en la Red se usan fundamentalmente dos métodos: navegar o buscar. Navegar significa seguir un recorrido hipertextual, lo que resulta bastante sencillo e intuitivo. Y buscar es expresar una necesidad de información a través de palabras claves que se combinan entre ellas, con ayuda de operadores y/o símbolos matemáticos y se forma una ecuación de interrogación. Una vez obtenidos los resultados de la primera búsqueda, estos se revisan y, dependiendo del grado de relevancia y pertinencia de las respuestas, se decide si hay necesidad de reformular la expresión, agregar nuevos términos, cambiar el idioma, utilizar sinónimos, etc., con el fin de refinar la operación de búsqueda.

Aplicar el método de buscar información en la Web implica conocer las características propias de los documentos existentes en este soporte, así como los principios básicos de la teoría de la recuperación de información, específicamente se trata de:

- Identificar palabras claves
- Seleccionar sinónimos y formas variantes en otros idiomas de las palabras claves
- Combinar sinónimos y palabras claves
- Elegir los instrumentos de búsqueda. Si se eligen BD se deben traducir los términos seleccionados al lenguaje documental (tesauros) para interrogar la BD.
- Formular la ecuación de búsqueda con la ayuda de los operadores booleanos (AND = *, OR = +, NOT = -), los posicionales (ADJ, NEAR, FAR, IN, WITHIN, AFTER, BEFORE) y los de exactitud, también llamados comodines como por ejemplo el truncado de palabras con el asterisco (*), cierre de interrogación (?) o signo de dólar (\$). Este último se usa para obtener resultados más amplios a partir de la raíz de las palabras.

En ese mismo orden, para lograr mayor eficiencia en la búsqueda y recuperación de información en la WWW, Cohen (2001) sugiere un enfoque basado en la pregunta, centrado en la búsqueda individual, que precisa de la curiosidad por descubrir la utilidad de las herramientas según los tipos específicos de búsquedas, dicho conocimiento proporciona la razón para su uso a posteriori. Aprender a identificar las necesidades de información y encontrar soluciones. La

idea no es enumerar una lista de características de cada herramienta, sino comprender la utilidad de cada una de ellas dentro del contexto de la búsqueda de información.

El enfoque de Cohen (2001) es justamente una estrategia de búsqueda que se traduce en:

1. Aplicar lo que ya se sabe acerca de los recursos de las bibliotecas para buscar herramientas en la web.

En Internet se reproduce la misma variedad de fuentes de información que se pueden encontrar en las bibliotecas, se publican: catálogos, BD, obras de referencias (diccionarios, enciclopedias, glosarios, directorios), libros, publicaciones periódicas y seriadas. Además de los nuevos tipos de documentos que surgen cada día con características propias de ese nuevo medio electrónico: Portales, *Weblogs*, *Webrings*, *Wikis*, etc.

2. Seleccionar la herramienta que puede ayudar a encontrar lo que se necesita.

Generalmente se piensa que una herramienta puede responder a todas las preguntas. Si después de probar una herramienta, los resultados no son los esperados, es preciso intentar con otros instrumentos de búsqueda diferentes. Vale recordar que cada uno de ellos sirven a diferentes propósitos, por ello cabe preguntarse cuales herramientas se adaptan o ajustan más a cada necesidad.

3. Definir la búsqueda de manera amplia.

Se refiere a identificar la necesidad de información. Lo cual va más allá de la construcción de una frase de búsqueda. Vale preguntarse si el tema es amplio, incluye sub-temas o si es un tópico dirigido hacia un concepto limitado o construido de múltiples conceptos. En algunos casos es probable que la búsqueda sea más precisa y directa si se hace en un sitio web específico. También la búsqueda podría limitarse a un conjunto de datos como objetivo o a una información que cambia de manera dinámica.

4. Descubrir y utilizar otras herramientas, no solo motores de búsquedas.

En la Web, además de los Motores de búsqueda hay otros tipos de herramientas para buscar información: Metabuscadore, Directorios, *Weblogs*, Portales, Agentes personales, BD de publicaciones electrónicas, Catálogos de bibliotecas y Web profundo o Internet invisible. Es necesario prepararse en el uso de cada una de ellas porque cada herramienta de manera individual devuelve resultados distintos que pueden ser más o menos pertinentes a la necesidad planteada.

5. Analizar las necesidades de información e identificar las herramientas que soportan la solución o respuesta de las mismas.

Una vez que se conoce la diversidad de herramientas de búsqueda en la web y se tiene claridad de cual es la necesidad de información, se está en capacidad de hacer corresponder las herramientas que permitirán obtener resultados más eficaces con la solución para cada una de las interrogantes que conforman dicha necesidad.

6. Hacer las búsquedas generales o de temas que necesiten exploración en los Directorios.

Estas fuentes de información presentan listados de temas jerárquicos e incluyen sub-tópicos que ayudan a familiarizarse con el enfoque temático e ir refinando la búsqueda. También incorporan metadata (datos acerca de datos) que sirven de punto de partida para la investigación.

7. Usar los motores de búsqueda tanto para aquellas búsquedas dirigidas a un objeto preciso, búsquedas ambiguas como para asuntos amplios.

Estos permiten hacer búsquedas desde las más precisas hasta las muy amplias porque los límites pueden ser puestos en uso a través de la sintaxis de la frase de búsqueda o por selecciones en un menú de una plantilla o formulario de búsqueda. En muchos casos se utilizan delimitadores de campo tales como: localización geográfica, última fecha de modificación, idioma, o tipo de archivo.

8. Asumir que las BD o archivos no textuales generalmente se encuentran en la web profunda.

Solo una parte del contenido disponible en la WWW, fundamentalmente páginas estáticas, es accesible por medio de los buscadores convencionales. La parte mayor se encuentra en BD de archivos no textuales: datos, gráficos, software, información que cambia de manera dinámica como los multimedia.

Los contenidos en la Red de redes pueden ser recuperados por distintas vías; muchas BD son recuperables solo desde sus propios sitios web y estos sitios son accesibles desde directorios y motores de búsqueda. Pero ya hay muchos motores de búsqueda que incorporan nuevos instrumentos para rastrear la web profunda como por ejemplo el *Google Scholar*⁴. Otros sitios se han dedicado a reunir una colección de enlaces a BD recuperables en la WWW, ese es el caso de *Webfile*⁵ o *Academic Info*⁶ y otros tales como *ProFusion*⁷, buscan el contenido de BD seleccionadas desde una sola y única interfase.

9. Evitar una clasificación rígida de las herramientas por sus funciones.

El ambiente digital es muy cambiante, con mucha frecuencia surgen actualizaciones o modificaciones de las herramientas, por ello no tiene sentido atribuirles un uso único y específico a cada una, más bien la selección debe basarse, en la naturaleza de la pregunta y tener presente el contexto al que pertenece.

10. Evitar el estrés ante la infoxicación⁸.

Según Cornella (2000) "la realidad es que tenemos mucha más información de la que podemos manejar"(¶ 3). Al mismo tiempo, la volatilidad del mundo digital no se detiene y es imposible llevar el control de todo lo que se produce y multiplica en este medio. Hay que aprender a seleccionar la información crítica⁹ de cada uno para estar en capacidad de discriminar lo que es útil de lo que no lo es.

⁴ <http://scholar.google.com>, busca artículos y textos producidos en las universidades, el mundo académico.

⁵ <http://webfile.com/>

⁶ <http://www.academicinfo.net/>

⁷ <http://www.profusion.com/index.htm>

⁸ Según Cornella (2000) infoxicación es el exceso de información

⁹ Aquella información de la cual no se puede prescindir, la que requiere una actualización permanente y sistemática y que se puede reunir y organizar bajo cinco temas primarios y cinco secundarios (Cornella, 2000).

El decálogo de Cohen (2001) es una guía que ayuda a mejorar la búsqueda y recuperación en Internet, pero a estas estrategias hay que sumar otras ideas como las propuestas de Sherman y Price (2001) para buscar en la web profunda:

- adoptar la mentalidad del cazador: mente activa, de acecho permanente, que aprovecha las ventajas del ambiente;
- usar los motores de búsqueda para localizar las puertas de entrada a los sitios ocultos y una vez ahí, interrogar las BD según los requerimientos propios del sistema;
- crear una colección propia de direcciones, enlaces, archivos y listas de recursos organizados para facilitar su uso en el futuro;
- localizar grupos o individuos en el ciberespacio que se han dedicado a reunir lo mejor de lo mejor en asuntos de un interés específico;
- crear un servicio propio de monitoreo web;
- participar en redes de colaboradores;
- usar senderos para entrar en la web invisible y revisar las ayudas para entender las características propias de cada herramienta y aprovechar al máximo sus potencialidades.

El otro elemento que no se puede descuidar durante el proceso de búsqueda y recuperación de información en Internet es conocer y aplicar criterios de evaluación para los documentos electrónicos, de modo que se esté en capacidad de distinguir la información válida y confiable dentro de ese inmenso mar de documentos e informaciones en el que abunda tanto contenido desechable. De acuerdo con Salazar (2005) “la clave está en no fiarse de la primera impresión y en contrastar siempre la información” (p.99).

Por ejemplo, en aras de comprobar la autoridad se deben localizar currículos, bibliografías, citaciones, u otras experiencias que den muestras de la rigurosidad del autor. También es importante observar los contenidos con sentido crítico para poder distinguir entre la información verdaderamente objetiva y la que pueda estar sesgada.

Sin ánimo de proponer una lista de recursos de información en línea, se ofrecen algunas pistas para identificar fuentes de información en Internet que se consideran de alta confiabilidad y utilidad práctica en el ámbito de la comunicación profesional, y con las que se puede iniciar la colección propia de enlaces y listas de recursos organizados a la que hacen referencia Sherman y Price (2001).

En primer lugar destacan los sitios web gubernamentales o fuentes oficiales e institucionales como los ministerios y los organismos autónomos. Estos sitios aparecen identificados en el dominio con el *.gov*¹⁰ y tienen un peso importante para el periodismo, no sólo por la confiabilidad de la fuente, sino también porque publican informaciones estadísticas y legislativas, que son muy relevantes en esta profesión marcada por la presión en el tiempo de respuesta. En Venezuela por ejemplo se encuentran los siguientes:

- Instituto Nacional de Estadística (INE)¹¹ ofrece información sobre censo de población, estadísticas municipales, fuerza de trabajo, comercio exterior,

¹⁰ Ver Resolución 240 del Ministerio de Ciencia y Tecnología en la que se establece el cambio a *.gob* antes del 12 de noviembre de 2006. <http://www.nic.ve/view/noticiaShow.php?id=7>

¹¹ <http://www.ine.gov.ve/ine/indexine.asp>

índice de desarrollo humano u otros. Además desde esta página web se pueden localizar enlaces de interés como servidores de información estadística por continentes.

- Instituto para la Defensa y Educación del Consumidor y del Usuario (INDECU)¹² ofrece estadísticas sobre denuncias, conciliaciones y sustanciaciones.
- Ministerios: Energía y Minas¹³, Finanzas¹⁴ Educación y Deportes¹⁵, Instituto Venezolano de Seguros Sociales¹⁶, Ministerio Público¹⁷ y otros
- Sistema Integrado de Indicadores sociales para Venezuela¹⁸
- Banco central de Venezuela¹⁹
- El Portal Gobierno en línea²⁰

Otras fuentes muy relevantes para el periodismo son las que ofrecen los organismos internacionales, entre ellos: La Comisión Económica para América Latina y El Caribe²¹ (CEPAL). Censo General de Población y Vivienda en Venezuela, XIII (2001)²² y cada una de las agencias que conforma la Organización de Naciones Unidas: UNESCO, PNUD, OPS/OMS, UNIFEM UNOJIN²³, UNODC²⁴ y otros.

Asimismo, vale la pena identificar sitios especializados tanto nacionales como internacionales y, de acuerdo con la heterogeneidad temática que abarquen los intereses específicos de cada profesional, se pueden agrupar por categorías temáticas, geográficas, cronológicas u otras.

Dentro de este tipo de recursos para Venezuela resultan de especial interés, entre otros a nivel internacional: CIA World Factbook²⁵, que recoge información demográfica, geográfica, económica y política de todos los países del mundo; la BD de Políticos de las Américas²⁶. Éste es un proyecto conjunto de Gobernabilidad Democrática del Centro de Estudios Latinoamericanos en Georgetown University y la Unidad para la Promoción de la Democracia de la Organización de Estados Americanos (OEA), que recopila información documental y estadística sobre los 35 países del continente, constituciones políticas, leyes electorales, información sobre el ejecutivo, legislativo y judicial de cada país. Además incluye listas de diarios electrónicos y direcciones electrónicas de funcionarios gubernamentales; La Red Europea de Documentación e Información

¹² <http://www.indecu.gov.ve/>

¹³ <http://www.mem.gov.ve/>

¹⁴ <http://www.mf.gov.ve/>

¹⁵ <https://www.me.gov.ve/>

¹⁶ <http://www.ivss.gov.ve/>

¹⁷ <http://www.fiscalia.gov.ve/>

¹⁸ <http://www.sisov.mpd.gov.ve/>

¹⁹ <http://200.74.218.51/>

²⁰ <http://www.gobiernoenlinea.ve/misc-view/index.pag>

²¹ <http://www.eclac.cl/estadisticas/>

²² <http://www.pnud.org.ve/censo2001/principal.htm>

²³ <http://www.uncjin.org/> Centro para la Prevención del Crimen Internacional

²⁴ <http://www.unodc.org/unodc/index.html> Oficina contra la Droga y el Delito de las NU

²⁵ <http://www.cia.gov/cia/publications/factbook/geos/ve.html>

²⁶ <http://pdba.georgetown.edu/>

sobre América Latina²⁷; vLex²⁸, la legislación en Internet; el sitio del Reportero investigador²⁹; Infoamérica³⁰, Portal de la Comunicación de la UAB³¹; Sala de Prensa³²; Biblioteca Universitaria de la Universidad de Málaga (UMA³³) y el Índice de Bibliotecas de Internet³⁴

A nivel nacional, entre los sitios que contienen valiosos recursos de información especializado destacan: Datanalisis³⁵, empresa de investigación de mercados especializada en las áreas de economía, estrategias, análisis de mercados y análisis de políticas públicas; la Cámara Venezolana de Comercio Electrónico (Cavecom-e³⁶), organización privada, independiente y sin fines de lucro que agrupa en su seno a las empresas y personas naturales que basan o apoyan sus actividades en los medios digitales y constituyen un gremio transversal a casi todos los sectores de la vida económica; la Revista Producto³⁷, especializada en publicidad y mercadeo; el Programa Venezolano de Educación Acción en Derechos Humanos (PROVEA³⁸); el Servicio Jesuita de Refugiados³⁹ y la Organización Nacional de Derechos Humanos de la Niñez y Adolescencia (CECODAP⁴⁰); Espacio público⁴¹, que trabaja en la promoción del derecho a la comunicación de los ciudadanos, la libertad de expresión y la práctica de una comunicación social de calidad y Venmedios⁴² que provee información estratégica del entorno mediático nacional.

Fuentes académicas en Venezuela

El Centro de Investigación de la Comunicación de la Universidad Católica Andrés Bello (CIC-UCAB) como institución académica asumió la responsabilidad de convertirse en un proveedor de contenido confiable, por ello trabaja en la creación y desarrollo de un sistema de información documental, orientado a satisfacer las necesidades de los usuarios de la comunidad académica (estudiantes, profesores, investigadores), sin excluir a otros interesados en el campo de la Comunicación social y la Cultura venezolana.

Este sistema, incorpora tanto BD referenciales (RECOM⁴³) como de texto completo (SVI⁴⁴), y permite la publicación de contenidos multimedia. Las BD han

²⁷ <http://www.redial-ceeib.net/portada.php>

²⁸ <http://www.vlex.com/contenidos/>

²⁹ <http://www.ire.org/esp/>

³⁰ <http://www.infoamerica.org/>

³¹ <http://www.portalcomunicacion.com/>

³² <http://www.saladeprensa.org/>

³³ <http://www.uma.es/Servicios/Biblioteca/RecursosCComunicacion.html>

³⁴ <http://lii.org/>

³⁵ <http://www.datanalisis.com/>

³⁶ http://www.cavecom-e.org.ve/bin_cavecome/main/templates/home.asp?seccID=1

³⁷ <http://www.producto.com.ve/>

³⁸ <http://www.derechos.org.ve/>

³⁹ <http://www.serviciojesuitaarefugiados-vzla.org/>

⁴⁰ <http://www.cecodap.org.ve/hfm/pronuncia/01.htm>

⁴¹ <http://www.espaciopublico.org>

⁴² <http://www.venmedios.com/entorno.php>

⁴³ <http://200.2.12.152/cic/portal/index.html> Red de información y documentación en Comunicación y Cultura

⁴⁴ <http://www.ucab.edu.ve/ucabnuevo/SVI> Salas Virtuales de Investigación

sido diseñadas bajo estructuras que contemplan la descripción de metadatos para facilitar la clasificación y descripción normalizada, que garantice la posterior recuperación de las publicaciones eficientemente. Las herramientas de recuperación de información utilizadas en dicho sistema posibilitan la realización de búsquedas libres, por palabras y avanzadas, haciendo uso de los operadores booleanos. Su fortaleza está en la calidad de los documentos de su biblioteca virtual y en la ergonomía de sus interfaces de acceso.

Portal RECOM

Por un lado, se encuentra el conjunto de BD referenciales (RECOM) accesible a través del sitio web de la UCAB, que ofrece información bibliohemerográfica, descripción de imágenes, audiovisuales y directorio de medios, con facilidades de búsqueda consolidada en todas las bases a la vez, o en cada una por separado. Con este producto, se pretende conformar una red nacional que pueda sumarse a las iniciativas regionales e internacionales existentes en temáticas afines, como es el caso de la Biblioteca Virtual para las Ciencias Sociales de CLACSO.

Salas Virtuales de Investigación

La digitalización de archivos de interés mundial y su incorporación a Internet es un reto que se han planteado varias bibliotecas e instituciones de distintos países. El objetivo es mejorar los contenidos disponibles en la Red y elaborar fondos electrónicos sobre patrimonios culturales.

La sistematización de este proceso, llevado a cabo por la Unesco, a través de su proyecto "Programa de Memoria del Mundo", y la garantía de que colectivos de distintas áreas trabajen en ello, significará en un futuro cercano la posibilidad de acceder a materiales relevantes que ayudarán a entender el pasado y contextualizar el futuro.

Si bien estos proyectos se iniciaron en el campo científico con el MIT (Instituto de Tecnología de Massachusetts), el cual decidió hacer público su acervo documental, actualmente se encuentran experiencias en el área humanística como el portal de la Biblioteca Miguel de Cervantes, el Proyecto Gutenberg (que se dedica a colocar en línea textos cuyo derecho de autor ya vencieron), y el Proyecto Bartlevy de la Universidad de Columbia, entre otras.

En este contexto, en Venezuela existe una preocupación por rescatar la memoria periodística nacional y permitir su acceso en línea. Específicamente el Centro de Investigación de la Comunicación de la Universidad Católica Andrés Bello desarrolló el proyecto de las Salas Virtuales de Investigación, las cuales ofrecen información relevante sobre un tema o personaje importante del periodismo venezolano.

Esta iniciativa surgió una década atrás cuando Sofía Ímber, periodista pionera de la televisión venezolana, ofreció a la UCAB los archivos de los

programas que presentó junto a Carlos Rangel en diferentes espacios televisivos desde 1969 hasta 1993.

La transcripción de las entrevistas de los programas “Buenos Días”, “Lo de Hoy” y “Solo con Sofía” constituían testimonios del debate político, económico, social y cultural del ámbito nacional e internacional, al contar con invitados de destacadas trayectorias. Ante el valor histórico que representaba este archivo, se decidió rescatarlo, conservarlo en formato digital y permitir su acceso en Internet.

Una vez que esta experiencia logró su objetivo, a partir de la coordinación de diferentes áreas que conjugaban el trabajo informático, documental y periodístico, se crearon otras salas que destacaban la labor de los periodistas venezolanos Miguel Otero Silva, Carmen Clemente Travieso, y Ramón J. Velásquez.

El interés de este proyecto por la comunidad periodística e investigadora generó, además, otros espacios dedicados a la Prensa de la emancipación, escrita entre 1808 y 1812, y a la Fotografía Venezolana, el cual cuenta con diversas colecciones de imágenes que retratan el país y su cultura desde la década de los 50.

En total son seis salas dedicadas al ámbito periodístico que buscan convertirse en referencia de material académico disponible en Internet. Una vez culminada la primera etapa de este proyecto, en el que se logró la publicación en línea de los archivos de texto completo e imágenes y la puesta en marcha del motor de búsqueda para recuperar dichos documentos, el interés ahora es promover la creación de comunidades de investigadores que utilicen la herramienta, y además intercambien sus indagaciones y conocimientos en ese espacio virtual.

El CIC ha considerado continuar con esta segunda fase, al apoyar el planteamiento de que “el verdadero problema no es la satisfacción de las necesidades de información preexistente, sino la considerable ampliación del campo de la información, es decir, la automatización, la organización, la sistematización de informaciones tradicionales y la creación de informaciones nuevas”, (Dominique Wolton, 2000, p.102).

La evolución del concepto de las Bibliotecas Virtuales son prueba de ello, ya que en la actualidad no se basan únicamente en la consulta de catálogos en línea, sino que además han incorporado en su estrategia la posibilidad de acceder a textos plenos e imágenes, y a nuevos ambientes de intercambio.

Desde el ámbito académico se considera que el desarrollo de mejores recursos que fomenten la investigación y el debate en línea fortalecerá la especialidad del “ciberperiodismo”, entendida tal como lo explican Díaz Noci y Salaverría (2003) como aquella que emplea el “ciberespacio” para “la investigación, la elaboración y, muy especialmente, la difusión de contenidos periodísticos”(p.17).

Otras Universidades del país también están trabajando en el desarrollo de recursos electrónicos de calidad como es el caso de la Universidad Central de Venezuela con la creación de la Biblioteca Virtual UCV⁴⁵ y Saber-ULA ⁴⁶.

⁴⁵ <http://www.sicht.ucv.ve:8080/bvirtual/>

⁴⁶ <http://www.saber.ula.ve/>

Referencias

- Bergman, M. (2001, Agosto). The Deep Web: Surfacing Hidden Value. *The Journal of Electronic Publishing*. 7 (1), Recuperado el 02 de junio de 2006 de, <http://www.press.umich.edu/jep/07-01/bergman.html>.
- Cervera, J. (2001, mayo 28). *La nueva economía IP*. Recuperado el 03 de mayo de 2006 de, <http://www.baquia.com/noticias.php?idnoticia=00017.20010528>
- Codina, L. (2000). *La documentación en los medios de comunicación: situación actual y perspectiva de futuro*. [cd-rom]. En: Cuadernos de Documentación Multimedia. Madrid: MULTIDOC. n.10. Número especial dedicado al Primer Congreso Universitario de Ciencias de la Documentación. ISSN 1575-9725.
- Cohen, L. B. (2001, noviembre). 10 Tips para la Enseñanza de cómo buscar en el WEB. *American Libraries Associations (ALA)*: Illinois, 32 (11). Traducción en versión libre del Inglés.
- Cornella, A. (2002). *Como sobrevivir a la infoxicación*. Transcripción de la conferencia del acto de entrega de títulos de los programas de Formación de Postgrado del año académico 1999-2000. Recuperado el 11 de noviembre de 2002 de, <http://www.uoc.edu/web/esp/articles/cornella/acornella.htm>
- De Oteyza, C. (2003). *La memoria periodística venezolana en Espacios virtuales de investigación*. Ponencia no publicada, presentada en la X Jornada Nacional sobre investigación y docencia en la ciencia de la historia.
- Garrido, N. (2006). *Estrategias y herramientas para la redacción de notas periodísticas transparentes*. Ponencia no publicada, presentada en el Taller de Periodismo de Transparencia, Caracas: Universidad Católica Andrés Bello.
- González, C. (2003, agosto). *Arquitectura de la información: diseño e implementación*. Artículo presentado en el Taller de Infotech, Lima.
- Gulli, A. y Signorini, A. (2005). *The indexable web is more than 11.5 billion pages*. [en línea]. Recuperado el 5 de junio de 2005 de, <http://www.cs.uiowa.edu/~assignori/web-size/size-indexable-web.pdf>
- Martínez, A. y Santamaría, L. (1993). *Manual de estilo*. Indianápolis, Indiana: Centro Técnico de la Sociedad Interamericana de Prensa.
- Martínez, M. (s. f.). *Proyecto Gutenberg, Alex y Gabriel en Internet*. Recuperado el 12 de abril de 2006 del sitio web de la Universidad de Viena, http://www.univie.ac.at/Schroedinger/1_97/memo10.htm

Price, G. y Sherman, C. (2001, Julio - Agosto). Exploring the Invisible Web. *Online*. 25 (4), 32-34.

Rubio, M. (2005). Nuevos tiempos para la documentación informativa en el periodismo digital: viejas y nuevas funciones del servicio de documentación digital. *Comunicación y sociedad*. 18 (1), 153-168.

Salazar, I. (2005). *Las profundidades de Internet*. Gijón (España): Ediciones Trea, S.L.

Toledo, L. (2003). Entrevista AMB Isidro F. Aguillo. El verdadero problema de los motores de búsqueda de Internet no es el ruido documental sino el silencio. *Revista d' Arxius i Biblioteques: Compactus*. (8), 52-53.

Tramullas, J. (2002). Recuperación de información en el World Wide Web. López, J. (coord.). *Manual de Ciencias de la Documentación*. Madrid: Pirámide, pp. 601-631.

Universidad Católica Andrés Bello. Centro de Investigación de la Comunicación. (2006). *Salas Virtuales de Investigación*. [Archivo de datos]. Caracas, Venezuela: Centro de Investigación de la Comunicación (CIC-UCAB). Disponible en el sitio web de la Universidad Católica Andrés Bello, www.ucab.edu.ve/ucabnuevo/SVI

Varlamoff, M. (2000). *Estudio Ifla/Unesco sobre digitalización y conservación*. Recuperado el 5 de abril de 2005 de, http://www.unesco.org/webworld/mdm/survey_index_es.html

Wolton, D. (2000). *Internet ¿y después?: una teoría crítica de los nuevos medios de comunicación*. Barcelona: Gedisa.

Modo de citar este trabajo según las normas APA

Calderín M. y Rojano, M. (2006). El periodista y la documentación digital: nuevo espacio para la investigación. En: M. Rojano (coord.), *Diez años del periodismo digital en Venezuela*. Caracas: Publicaciones UCAB, pp. 273-292