

Análisis de la prensa cubana a partir de los resultados en buscadores

Víctor Angel Fernández

Índice

1	Introducción	1
2	Conclusiones	5
3	Anexo	5
4	Bibliografía	7

1 Introducción

El presente trabajo, forma parte de un conjunto de análisis sobre la presencia de la prensa cubana en Internet, los cuales incluyen el desarrollo de Internet en Cuba y la presencia de la prensa en esta Red de Redes. Pasan obviamente por algunos aspectos de la historia del desarrollo tecnológico en la prensa cubana, teniendo en cuenta que muchas veces cuando se habla de la entrada de nuevas tecnologías, sólo se piensa en las computadoras y se olvida desde el linotipo, hasta la concepción de un medio de prensa como gran empresa.

No se ha dejado pasar por alto el cómo “escribir” para Internet, en tanto referencia a la presencia de un medio en este entorno, con sus leyes y limitaciones. Por último, los trabajos llegan a proponer medios de evaluación y análisis de la prensa cubana en Internet, no sólo los tradicionales que se conocen, como pueden ser Web Site Garage y Hitbox, que ofrecen sus conocidos diagnósticos, sino

ir mucho más allá e introducirse en métodos de evaluación matemática, no frecuentes en el trabajo de prensa pero sí, muy conocidos en el trabajo de Bibliotecología y Biblioteconomía, en su más amplia concepción del término, como son los análisis informétricos y cuantitativos, que desde hace algún tiempo se están tratando de llevar a la evaluación de sitios web y han comenzado a acuñar el término “webometric” en inglés o sus variantes “webometría” o “webmetría” en español, sin dejar pasar algunos autores que simplemente optan por llamarle cibermetría a las referidas aplicaciones siempre que tengan que ver con Internet.

Aunque el horizonte de medios de prensa cubanos con presencia en Internet ya se acerca al centenar, se ha limitado la muestra a los órganos de circulación provincial, incluido Victoria, de la Isla de la Juventud, no obstante conocerse la inconsistencia de los términos municipal, provincial o nacional luego de tener presencia en Internet.

Es sabida la característica de Altavista como buscador, donde su algoritmo de clasificación o de ubicación, es casi unánimemente cuestionado por los autores dedicados al tema de análisis de máquinas de búsqueda y directorios, llegando en algunos casos al

extremo de desautorizarlo como base de análisis.

De igual forma, los autores consultados coinciden en señalar a Altavista, con sus características favorables y negativas, como el que mejor responde a solicitudes relacionadas con el tema de interés del presente trabajo, que es la evaluación matemática de la calidad de determinados sitios, sobre todo si se tienen en cuenta algunas posibilidades que sólo este buscador utiliza. Tal es el caso relacionado con la respuesta de URL, dominios y sobre todo con la combinación de los mismos con la búsqueda por frases o términos, lo cual incluso Google ni remotamente ofrece, y es en nuestro entender, hoy por hoy de acuerdo con las estadísticas, el más usado de los buscadores y el innovador en la ubicación del ranking debido a su análisis de “link popularity” (popularidad de enlaces).

No obstante lo planteado en el párrafo anterior, se realizaron pruebas con otros buscadores, probando en ellos las diferentes variantes necesarias para llevar a cabo el análisis previsto y los resultados confirmaron lo expresado por la bibliografía consultada y se ha utilizado Altavista en el desarrollo de la presente investigación. Aunque las opiniones sobre ranking, pueden ser confirmadas, esto no hace, ni con mucho a este buscador un elemento a despreciar, sobre todo si se tiene en cuenta que recibe unos 63 millones de visitas diarias y que en él se realizan más de 50 millones de búsquedas también diariamente.

Como puede apreciarse en el Anexo, estas son las posibilidades de búsqueda que ofrece Altavista y que se utilizaron para este trabajo, incluso la combinación de algunas de ellas, son las que potencian realmente el mismo.

Por ejemplo, en Google se pueden obtener los resultados de la búsqueda por palabras claves, ya sea de forma separada, con los operadores booleanos AND y OR o por medio de frases y combinaciones. De igual forma, permite buscar con la opción *link:www.sitio-buscado.cu*, pero no permite la combinación de esta opción con las palabras claves. Asimismo en los casos de que se solicite una búsqueda que incluya la partícula “www” frente al nombre del sitio, Google automáticamente elimina la misma por considerarla muy repetitiva, pero a su vez este elemento puede contaminar los resultados y, de hecho, alterar los mismos.

Teniendo en cuenta todo lo expresado hasta aquí, se ha partido de la base de análisis siguiente:

Ante una solicitud de información relacionada con el tema por el cual un sitio debe responder fundamentalmente, qué impacto tiene en la referida máquina de búsqueda. O sea, para el caso particular de la prensa, se seleccionaron los temas por los cuales debe responder la prensa cubana ante una solicitud de información en Internet. La base de análisis de términos utilizados fue la ofrecida por el sistema de estadísticas Stat Live, por el cual se conoce el tráfico de los sitios hospedados en Cubaweb, con la ventaja importante de que su acceso es totalmente abierto, o sea, no es necesario poseer login y password para conocer los resultados. Este sistema ofrece la opción de conocer las palabras claves más utilizadas a la hora de solicitar los referidos sitios.

A saber, estos términos de búsqueda son: (Entre paréntesis la cantidad de sitios que responden en Altavista a la búsqueda solicitada y la sintaxis de búsqueda utilizada en los casos de combinaciones).

- Cuba – (1 061 815)
- Prensa – (1 138 958)
- News – (189 695 241)

En este caso se utilizó el término en inglés con independencia del idioma en que esté el medio de prensa, pues se parte de la base de que el visitante teclee el término primero en su idioma, con independencia del resultado que pueda obtener y a partir de ahí establecerá la correspondiente discriminación.

- Noticias – (1 138 958)
- Cuba y Prensa – (69 819 – cuba AND prensa)

Sitios que responden a ambos términos al mismo tiempo y ello permite establecer mayores especificidades sobre la búsqueda, según puede apreciarse en los resultados. De acuerdo con las definiciones de Altavista, no existe ninguna diferencia entre los términos utilizados con mayúsculas o con minúsculas. De hecho tendrá el mismo resultado informativo, Informativo o InFoRmAtIvo, además de cualquiera de sus variantes.

- Prensa cubana – (914 584 – prensa cubana)

La máquina de búsqueda responde ante la solicitud de uno cualquiera de los dos términos o los dos al mismo tiempo.

- “Prensa cubana” (821 – “prensa cubana”)

La máquina de búsqueda responde a la frase compuesta y ello da mayor especificidad en la búsqueda.

- Cuba y News (569 160 – cuba AND news)

Igual al caso de combinar dos términos ya explicado

- Cuba y Noticia (12 829 – cuba AND noticia)

Otro caso de combinación de dos términos

en la misma búsqueda, con el objetivo de obtener mayor especificidad.

- Cuba y (Prensa o News o Noticia) (1 255 914 – cuba AND (prensa OR news OR noticia))

Se delimita el horizonte de búsqueda, al hacer más específica la misma, ya que a esta solicitud, responderán los sitios que incluyan al término Cuba, combinado al mismo tiempo con sólo alguno o todos de los términos entre paréntesis. El resultado es similar a solicitar Cuba y Prensa o Cuba y News o Cuba y Noticia

- Cuba y (Prensa o News o Noticia o Press) (1 860 396 – cuba AND (prensa OR news OR noticia OR press))

Similar explicación a la anterior, con la ampliación de otro término.

- Cuba y (Prensa o News o Noticia o Press) y Hospedaje=www.sitio.cu (cuba AND (prensa OR news OR noticia OR press) AND host:www.sitio-deseado.cu)

Similar al caso anterior, pero con la característica de buscar cómo responde cada órgano de prensa a estos términos, pues Altavista nos permite esa capacidad especial de búsqueda. La opción introducida (host:www.sitio-deseado.cu) utilizada independientemente, ofrece como respuesta cuántas veces está incluido en el buscador un sitio determinado, no los enlaces que se hagan a él. Los resultados de la respuesta, se expresan en la tabla adjunta, pues en esta oportunidad varían con cada medio de prensa

- Cuba y (Prensa o News o Noticia o Press) y Dominio=CU (118 – cuba AND (prensa OR news OR noticia OR press) AND domain:.cu)

Aquí la especificidad se hace todavía mayor, pues al caso anterior se le adiciona la solicitud de que la búsqueda sea referida

solamente a los sitios con dominio .CU, o sea, todos aquellos sitios generados en Cuba. Esta es otra de las opciones que ofrece Altavista, al permitir circunscribir la búsqueda, cualquiera que esta sea, a los sitios de un dominio determinado. Los resultados de la búsqueda independiente domain:.cu, o sea la cantidad de sitios registrados por el NIC cubano y que tienen presencia en Altavista, son 57 271.

Se ha obviado la capacidad de cada órgano de prensa para su respuesta en términos específicos, o sea, medios del turismo, la cultura, deportes u otras ramas que pueden utilizar estas consideraciones y sumarles sus particularidades, pero que en lo fundamental son “medios de prensa creados desde Cuba”, que sería aproximadamente la respuesta al último caso. Si bien es cierto que algunos (realmente pocos) órganos de prensa cubanos responden a diferentes dominios, como puede ser ORG o COM, pensando en los casos de Prensa Latina o el caso del Bazar de la Prensa Cubana, LatinCuba.com, su expresión es mínima y no influye de ninguna manera en la muestra.

Como puede apreciarse en los resultados que se ofrecen en cada caso, las cantidades son extremadamente grandes, del orden de los millones o los cientos de millones (caso particular de la respuesta a la solicitud PRENSA), por lo cual fue imposible comprobar los resultados caso por caso. Sólo en la última opción, cuyo resultado es 118, fue posible analizar cada referencia a quien pertenecía.

Lo primero que se hizo fue comprobar que el resultado era cierto, o sea, que eran sitios registrados en Cuba. A continuación se fue abriendo (accediendo) a cada uno de ellos y con la opción “*búsqueda sólo en esta*

página” que ofrece el browser Internet Explorer, se comprobó que en todos los casos los términos buscados estaban presentes en alguna de las secciones componentes de la página, ya sea TITLE, Meta Tags, o BODY.

Un elemento interesante fue el comprobar la cantidad de sitios, que no obstante no ser de un órgano de prensa propiamente dicho como pueden ser los portales de Infocom, Islagrande y Citmatel, aparecían incluidos presentando noticias o informaciones en su más amplio sentido. Es importante destacar que lo anterior es sólo una apreciación de los resultados de la búsqueda y no un punto de vista sobre el discutido concepto de “*Internet como medio de prensa*”, lo cual merece un trabajo por sí mismo.

El resultado de cuáles de esos 118 sitios o páginas eran realmente de cualquiera de los medios de prensa cubanos, fue sólo de 30, quedando excluidos de esta lista los correspondientes a Juventud Rebelde, Venceremos (Guantánamo), Vanguardia (Villa Clara), Girón (Matanzas), Notinet (Radio Reloj), Guerrillero (Pinar del Río), Invasor (Ciego de Avila), La Demajagua (Granma) y Victoria (Isla de la Juventud), pertenecientes a la muestra de órganos nacionales y provinciales, utilizados como base del trabajo de investigación.

Las respuestas a esta situación pueden ser varias pero en un análisis totalmente aleatorio de algunos de los sitios que no están incluidos, el primer aspecto que salta a la vista es la ausencia total de Metatags, como es el caso de Juventud Rebelde, Girón o Venceremos y en los demás casos, su presencia es mínima o no correctamente diseñada.

De igual forma, casi todos los órganos de prensa cubanos, utilizan muy poco las capacidades del título (sección TITLE del htm),

donde aparecen dos, tres o cuatro palabras promedio, no obstante conocerse que los sistemas internacionales de análisis, como pueden ser Web Site Garage o Hitbox, por sólo citar dos de los más conocidos, proponen una media de hasta 15 palabras. Tampoco incluyen en el título las palabras buscadas, o sea, Cuba, prensa, noticia, u otras similares. Y este es el primer elemento por el que evalúan casi todos los buscadores.

2 Conclusiones

Luego del desarrollo en este año 2001, donde ya en estos momentos se alcanza el centenar de órganos con presencia en Internet, toda la prensa provincial que se edita en papel, incluida la de la Isla de la Juventud tiene su sitio con diferentes niveles de desarrollo y profundidad del tratamiento de los temas. Por esa razón se escogieron estos sitios, sumados a los de los órganos de circulación nacional, como muestra de la investigación.

Se utilizó el buscador Altavista debido a sus características especiales de análisis, las cuales ofrecen posibilidades que no están presentes en el resto de los buscadores.

Asimismo se realizó una investigación previa donde se seleccionaron los términos más comunes por los que se busca la prensa cubana en Internet y esa fue la base de análisis.

Los resultados obtenidos, reflejan una mínima presencia de los sitios de la prensa cubana ante las referidas solicitudes de búsqueda. En ella saltan primero a la vista errores en la programación de las páginas que conforman los sitios, como es el caso de la composición de los títulos o la pobre composición de los Meta-Tags y también la ausen-

cia de los mismos en algunos de los medios de prensa.

De igual forma, una posible razón puede estar dada por un insuficiente trabajo de promoción de los sitios, debido al probable desconocimiento de las metodologías utilizadas para inscribir sitios en los buscadores, mantenerse en los mismos y después escalar posiciones que, según dicen todos los autores consultados, sólo cuando se está entre las 10 primeras respuestas (generalmente las que aparecen en la primera página de resultados de cualquier máquina de búsqueda), es que se posee un real posicionamiento en el referido lugar.

3 Anexo

Descripción de la Hoja de Búsqueda de Altavista

And

Busca documentos que contienen todas las palabras especificadas o las combinaciones de frases. La solicitud agua AND tierra, devolverá los documentos que respondan a los dos términos al mismo tiempo.

Or

Encuentra los documentos que contenga al menos uno de los términos especificados. En esta caso, la solicitud agua OR tierra responde con los documentos que contengan los dos o uno cualquiera de los términos.

And not

Excluye los documentos que contienen una palabra o una frase en específico. Siguiendo con el ejemplo, agua AND NOT tierra, devolverá las respuestas que se refieran a “agua” y al mismo tiempo no tengan

referencia al término “tierra”.

Near

Encuentra los documentos que contengan las palabras especificadas con una distancia en el término no mayor de 10 palabras entre ellas. La solicitud agua NEAR tierra, responderá, por ejemplo, con un documento de este contenido agua palabra1 palabra2 palabra3 tierra, pero, de igual forma no responderá con un documento de este tipo: agua palabra1 palabra2 ... palabra10 palabra11 tierra.

()

Los paréntesis agrupan elementos complejos de las búsquedas booleanas. Por ejemplo, (agua AND tierra) OR (barcos AND autos).

Anchor:texto

Encuentra las páginas que contienen la palabra especificada, pero sólo cuando está formando parte de un enlace o vínculo. Por ejemplo, anchor:agua responderá con la página que tenga un enlace de este tipo: <a href:http://www.unsitio.com>agua. Es importante no poner espacios después de los dos puntos (:).

Applet:class

Encontrará las páginas que contengan un Applet de java con ese nombre.

Domain:nombrededominio

Su respuesta, son las páginas de sitios dentro de ese dominio. Por ejemplo, domain:cu, devuelve todos los sitios o páginas procesados por Altavista que provengan del dominio CU.

Host:nombredelhost

Busca páginas de un hospedaje determinado.

Image:fichero

En este caso, la respuesta serán las páginas que contengan un nombre de imagen determinado. El nombre se refiere al fichero de las imágenes y no al valor “Alt”, que a veces se pone en las mismas.

Like:URLtexto

Aquí las respuestas son páginas cuya URL tenga alguna relación con la que se ha solicitado.

Link:URLtexto

Esta opción busca todas las páginas con enlaces a la URL especificada.

Text:texto

Busca las páginas que contienen el texto especificado, pero que no esté ubicado en un enlace, un Metatag o una imagen.

Title:texto

Busca páginas que contiene el texto especificado como parte de la sección TITLE del código htm.

Url:texto

En este caso las páginas buscadas son las que tienen en su URL la palabra especificada.

4 Bibliografía

<http://www.websitegarage.com>

Lugar para analizar características de los sitios.

<http://www.hitbox.com>

Lugar similar al anterior, que ofrece además otras herramientas gratuitas.

<http://www.altavista.com>

Repertorio de búsqueda utilizado para la investigación.

<http://www.google.com>

Repertorio de búsqueda

<http://www.alltheweb.com>

Repertorio de búsqueda

<http://www.yahoo.com>

Repertorio de búsqueda

<http://www.cubaweb.cu/stats>

Sitio de las estadísticas del portal cubano. Con esta sintaxis, o sea, nombre del sitio hospedado en Cubaweb, un slash "/" y la palabra "stats", pueden revisarse las estadísticas del sitio en cuestión

Es parte de la bibliografía, la tabla adjunta con las direcciones de los sitios de la prensa utilizados para el presente trabajo.

Granma Diario,

www.granma.cubaweb.cu

Granma Internacional,

www.granma.cu

Juventud Rebelde,

www.jrebelde.cubaweb.cu

www.bocc.ubi.pt

Trabajadores,

www.trabajadores.cubaweb.cu

Agencia de Información Nacional,

www.ain.cubaweb.cu

NotiNet de Cuba,

www.nnc.cubaweb.cu

Guerrillero (Pinar del Río),

www.guerrillero.co.cu

El Habanero (La Habana),

www.elhabanero.cubaweb.cu

Tribuna de La Habana (Ciudad de La Habana),

www.tribuna.islagrande.cu

Girón (Matanzas),

www.giron.cubaweb.cu

Vanguardia (Villa Clara),

www.vanguardia.cubaweb.cu

5 de Septiembre (Cienfuegos),

www.5septiembre.cu

Escambray (Sancti Spíritus),

www.escambray.islagrande.cu

Invasor (Ciego de Avila),

www.invasor.islagrande.cu

Adelante (Camagüey),

www.cmw.inf.cu/adelante

26 (Las Tunas),

www.periodico26.cu

La Demajagua (Granma),

www.lademajagua.co.cu

Ahora (Holguín),

www.ahora.cu

Sierra Maestra (Santiago de Cuba),
www.sierramaestra.cu

Venceremos (Guantánamo),
www.venceremos.cubaweb.cu

Victoria (Isla de la Juventud),
www.islagrande.victoria.cu

En los casos de los sitios de la prensa que tienen dos direcciones (una por Cubaweb y otra por Infocom, esta última como “co.cu”) se revisaron ambas, pero en el segundo caso, no surgió ninguna respuesta de relevancia que variara o que de cualquier forma influyera en los resultados de la muestra.

Tabla

Granma – 603
Trabajadores – 1
Juventud Rebelde – 0
Granma Internacional – 2210
Agencia de Información Nacional – 103
Notinet del Cubaweb (NNC) – 111
Guerrillero (Pinar del Río) – 0
El Habanero (La Habana) – 1
Tribuna de La Habana – 2
Girón (Matanzas) – 0
Vanguardia (Villa Clara) – 0
5 de Septiembre (Cienfuegos) – 1
Escambray (Sancti Spíritus) – 58
Invasor (Ciego de Avila) – 0
Adelante (Camagüey) – 0
26 (Las Tunas) – 1
Ahora (Holguín) – 2
La Demajagua (Granma) – 1
Sierra Maestra (Santiago de Cuba) – 1
Venceremos (Guantánamo) – 0
Victoria (Isla de la Juventud) – 0