



Universidad de las Ciencias  
Informáticas

**Universidad de las Ciencias Informáticas.**

**Facultad 10**

**Título:** “Los indicadores webmétricos y su influencia en la calidad de los sitios web.”

**Trabajo de Diploma para optar por el título de:**

*Ingeniero en Ciencias Informáticas*

**Autor(es):**

*Danny Fernández Herrera  
Fidel Enrique Blanco Sánchez*

**Tutor(es):**

*Ing. Yonny Mondelo Hernández.  
Lic. Paula Yuleni Hechevarría Leyva*



*Uno de los factores que determinan la gran diferencia entre el mundo desarrollado y el mundo en desarrollo (...) un elemento determinante de esa enorme diferencia, está en la ciencia.*

A handwritten signature in black ink, which appears to be 'Fidel Castro Ruz'.

*Fidel Castro Ruz*

## Declaración de Autoría

Declaran ser los únicos autores del trabajo titulado:

Los indicadores webmétricos y su influencia en la calidad de los sitios web.

Y reconocen a la Universidad de las Ciencias Informáticas los derechos patrimoniales de la misma, con carácter exclusivo.

Para que así conste firman la presente a los \_\_\_\_ días del mes de \_\_\_\_ del año 2010.

\_\_\_\_\_  
*Firma del Autor*  
Danny Fernández Herrera

\_\_\_\_\_  
*Firma del Autor*  
Fidel E. Blanco Sánchez

\_\_\_\_\_  
*Firma del Tutor*  
Yonny Mondelo Hernández

\_\_\_\_\_  
*Firma del Tutor*  
Paula Yuleni Hechevarría Leyva

## Resumen

La siguiente investigación tiene como título: *“Los indicadores webmétricos y su influencia en la calidad de los sitios web”*. Ha sido realizada con el propósito de mejorar la calidad de los sitios web en la Universidad de las Ciencias Informáticas (UCI). Para ello, se examina la evolución de dos principales disciplinas métricas de la información: la Cibermetría y la Webmetría, y se abordan algunas de sus características. Se estudian y analizan los principales indicadores webmétricos que deben tenerse en cuenta en la optimización de un sitio web, seleccionando para el estudio los que más valor le dan a la calidad de estos sitios. A partir de dichos indicadores, se analiza su comportamiento en un conjunto de sitios web presentes en Internet y en la Web de la UCI. Los resultados obtenidos manifiestan aquellos aspectos que deben ser mejorados para la completa calidad del sitio web. Finalmente se presenta una estrategia metodológica que contribuya a mejorar la calidad de los sitios web en la UCI.

# Índice general

<b>DECLARACIÓN DE AUTORÍA .....</b>	<b>II</b>
<b>RESUMEN.....</b>	<b>III</b>
<b>ÍNDICE GENERAL .....</b>	<b>IV</b>
<b>INTRODUCCIÓN.....</b>	<b>1</b>
<b>CAPÍTULO 1. INDICADORES WEBMÉTRICOS. ASPECTOS TEÓRICOS GENERALES. ....</b>	<b>7</b>
1.1. CIBERMETRÍA Y WEBMETRÍA. SURGIMIENTO, CONCEPTO Y EVOLUCIÓN. ....	7
1.1.1. Aplicaciones de la Cibermetría y la Webmetría. [5].....	9
1.2. INDICADORES CIBERMÉTRICOS. ....	11
1.3. CONCEPTOS RELACIONADOS CON LA CALIDAD.....	13
1.4. MOTORES DE BÚSQUEDAS O BUSCADORES. SU FUNCIONAMIENTO. ....	15
1.4.1. Funcionamiento de los motores de búsquedas. Relación con los sitios web.....	16
1.5. HERRAMIENTAS WEBMÉTRICAS. ....	17
1.6. ESTUDIOS WEBMÉTRICOS INTERNACIONALES Y NACIONALES. ....	22
<b>CAPÍTULO 2. SELECCIÓN DE LOS INDICADORES Y HERRAMIENTAS WEBMÉTRICAS PARA EVALUAR LA CALIDAD DE LOS SITIOS WEB.....</b>	<b>25</b>
2.1. INDICADORES WEBMÉTRICOS EMPLEADOS EN LOS ANÁLISIS MÉTRICOS.....	25
2.1.1. Indicadores webmétricos más empleados para la evaluación de los sitios web. ....	26
2.1.2. Indicadores webmétricos relevantes para estimar la buena elaboración de un sitio web. ....	28
2.2. HERRAMIENTAS INFORMÁTICAS EMPLEADAS PARA DESARROLLAR ESTUDIOS WEBMÉTRICOS.....	31
2.2.1. Herramientas Libres.....	33
2.2.2. Herramientas Online.....	34
2.3. SELECCIÓN DE INDICADORES WEBMÉTRICOS Y HERRAMIENTAS INFORMÁTICAS A UTILIZAR PARA EVALUAR LA CALIDAD DE LOS SITIOS WEB.....	35
2.3.1. Indicadores webmétricos seleccionados para el estudio.....	36
2.3.2. Herramientas informáticas seleccionadas para el estudio.....	39

<b>CAPÍTULO 3: ESTRATEGIA METODOLÓGICA PARA MEJORAR LA CALIDAD DE LOS SITIOS WEB EN LA UCI.....</b>	<b>43</b>
3.1.    FUNDAMENTACIÓN DE LA ESTRATEGIA. ....	43
3.1.1.    Componentes metodológicos de la estrategia.....	44
3.1.2.    Métodos y procedimientos utilizados para la elaboración de la estrategia: .....	45
3.1.3.    Establecimiento de principios o criterios de evaluación.....	47
3.2.    ANÁLISIS DE LOS SITIOS WEB REPRESENTATIVOS DE: LAS UNIVERSIDADES LATINOAMERICANAS SEGÚN EL RANKING WEB DE UNIVERSIDADES DEL MUNDO Y LOS SITIOS WEB DE LA UCI.....	47
3.2.1.    Evaluación de los sitios web correspondientes a las Universidades latinoamericanas. ....	47
3.2.2.    Análisis de los sitios web de la UCI.....	56
3.3.    ACCIONES PARA MEJORAR LA CALIDAD DE LOS SITIOS WEB DE LA UCI.....	62
<b>CONCLUSIONES GENERALES .....</b>	<b>68</b>
<b>RECOMENDACIONES.....</b>	<b>69</b>
<b>REFERENCIAS BIBLIOGRÁFICAS.....</b>	<b>70</b>
<b>BIBLIOGRAFÍA.....</b>	<b>73</b>

## Introducción

En la actualidad, el desarrollo de las Tecnologías de la Información y las Comunicaciones (TIC's), permite ampliar las posibilidades del progreso de la sociedad y con ello la adquisición de conocimiento útil, logrando satisfacer necesidades y apetencias profesionales. Como parte de esas tecnologías, se encuentra Internet, el cual es un medio interactivo que ofrece oportunidades para la comunicación y el intercambio de la información entre las personas. Sucede, que este universo digital está evolucionando muy rápidamente con un gran incremento de contenidos, lo cual provoca a cualquier persona la dificultad de tener un conocimiento actualizado en determinada esfera del saber.

Por lo tanto, este volumen de información disponible en Internet, requiere del uso de estrategias que permitan su aprovechamiento eficiente. Los sitios web constituyen una herramienta informática muy importante para procesar y presentar la información, logrando de esta forma que la misma sea utilizada eficazmente. La presencia y visibilidad en Internet de este recurso tecnológico, se vuelve sumamente esencial para lograr con éxito sus objetivos y obtener resultados concretos. Pero eso no lo es todo, la calidad de las páginas web es de vital importancia para alcanzar el triunfo, ya que un sitio web complejo y desorganizado, que no contenga profesionalidad en el diseño y no logre satisfacer las necesidades y expectativas de los usuarios de Internet, promueve un descenso de su calidad.

Una solución a los problemas que enfrenta la sociedad de la información digital, tales como: el crecimiento continuo a un ritmo exaltado de la información y el impacto de la calidad de los sitios web; la brindan los estudios métricos, apoyándose en el uso de los indicadores webmétricos, pues estos permiten organizar e interpretar adecuadamente toda la documentación del entorno digital. También proporcionan información cualitativa y cuantitativa que puede contribuir a mejorar la calidad, elevar el posicionamiento y por ende la visibilidad de una publicación o sitio web. De igual modo, posibilitan valorar el estado actual de las publicaciones y establecer pronósticos sobre las tendencias futuras para la toma de decisiones. De ahí que es oportuno el estudio de los indicadores webmétricos y denotar de esta manera cómo influyen sobre la calidad de los sitios web.

Estos estudios webmétricos han posibilitado que países como España, Brasil, Cuba y Perú, por solo nombrar algunos, se interesen en elevar su calidad en la Web y así contribuyan al aporte de nuevas ideas referente a la Webmetría. En Cuba, Ministerios de la Educación Superior y la Salud sobresalen en dichos

estudios, aportando a través de las revistas científicas u otros medios de información, los resultados de las investigaciones de profesionales de la información y la comunicación. Un ejemplo lo constituye la revista del Centro Nacional de Información de Ciencias Médicas de Cuba (ACIMED) la cual ha publicado entre otros artículos uno que lleva por título: “Indicadores cibernéticos: ¿Nuevas propuestas para medir la información en el entorno digital?” [1], abarcando de esta manera las disciplinas métricas de la información.

Los indicadores webmétricos usados para caracterizar sitios web constituyen un factor clave que determinan la calidad de las páginas web, entre otros aspectos. Actualmente, en la Universidad de las Ciencias Informáticas, no se valoran los criterios por autores reconocidos para la confección de un esquema, generalmente aceptado, que guíe la usabilidad de estos indicadores; se carece de un conocimiento sólido acerca del tema y se dificulta por parte de la mayoría de las personas de la institución, la mejor comprensión sobre la verdadera calidad de una o más páginas web.

Estos aspectos provocan que, aunque la UCI cuenta con el acceso a Internet necesarios para abastecer de conocimiento suficiente a los webmáster y diseñadores de los sitios web; dispone de los recursos tecnológicos y además del capital humano indispensables para la elaboración de estos sitios con calidad, en esta institución, se elaboren la mayoría de los sitios web, sin analizar a grandes rasgos, como influyen estos indicadores webmétricos sobre la calidad de los sitios en cuestión.

Con lo expresado anteriormente se evidencia por qué se hace necesario realizar el presente estudio, y de esta manera arribar correctamente a los mejores resultados.

Por lo que, a raíz de esta **situación problemática** se ubica el **problema científico** en: ¿Cómo mejorar la calidad de los sitios web en la UCI?

Siendo **objetivo de la investigación**: Elaborar una estrategia metodológica para mejorar la calidad de los sitios web en la UCI.

Como posible respuesta al problema científico, **o idea a defender**, se considera: Una estrategia metodológica para mejorar la calidad de los sitios web en la Universidad de las Ciencias Informáticas.

Determinando los siguientes **objetivos específicos**:



1. Definir los principales indicadores y herramientas webmétricas utilizadas para caracterizar los sitios web.
2. Analizar la influencia que tienen los diferentes indicadores webmétricos en la calidad de los sitios web.
3. Proponer una estrategia metodológica que ayude a mejorar paulatinamente la calidad de los sitios web de la Universidad de las Ciencias Informáticas.

En correspondencia con el problema científico y el objetivo planteado se especificaron las siguientes **preguntas científicas**:

1. ¿Qué indicadores webmétricos son los más usados para caracterizar los sitios web?
2. ¿Cómo influyen los indicadores webmétricos en la calidad de los sitios web?
3. ¿Qué nivel de importancia establecen los buscadores web a los distintos indicadores webmétricos?

El **objeto de estudio** son los indicadores cibernmétricos.

El **campo de acción** está centrado en los indicadores webmétricos.

Para lograr el objetivo propuesto, se plantean las siguientes **tareas de investigación**:

1. Fundamentar los aspectos teórico-metodológicos que sustentan los indicadores webmétricos.
2. Investigar los distintos indicadores webmétricos que se utilizan para caracterizar los sitios web.
3. Aplicar las herramientas informáticas que se emplean para realizar estudios webmétricos o mediciones específicas de determinados indicadores de la Web.
4. Diagnosticar la calidad actual de los sitios web en la UCI.
5. Establecer una estrategia metodológica que ayude a mejorar paulatinamente la calidad de los sitios web en la UCI.

Durante el proceso de la investigación se emplearon diferentes métodos empíricos y teóricos que estuvieron determinados por el objetivo y las tareas a desarrollar, que a continuación se reseñan:

## **Métodos Teóricos.**

- ✓ Histórico-lógico.

Se analiza la evolución y desarrollo del objeto de estudio así como su campo de acción, a partir de

estudios webmétricos llevados a cabo durante los últimos años a nivel nacional e internacional, además de investigaciones científicas sobre los indicadores webmétricos, mediante la cual se establece una base fundamental de conocimiento que ayuda a la comprensión del tema.

✓ Analítico-sintético.

Se realiza un análisis general sobre las teorías y documentos existentes relacionados con los diferentes indicadores webmétricos, con el fin de conocer más del objeto de estudio y su campo de acción, lo cual permite: explicar, hacer analogías y comprender mejor su comportamiento. De esta manera se extraen los indicadores más relevantes relacionados a la calidad de los sitios o páginas web.

✓ Inductivo-deductivo.

Se analizan casos particulares de los sitios web mediante el uso de los buscadores, que presenten en su diseño, la aplicación de indicadores webmétricos. De esta forma se generaliza el comportamiento de estos indicadores en dichos sitios web, para dar a conocer cómo se deben aplicar en la determinación de una adecuada calidad web en los mismos.

### **Métodos empíricos.**

✓ Observación.

Se hace un registro visual durante todo el proceso de detección de los indicadores webmétricos en determinados sitios web, de forma tal que se pueda recoger la información de los acontecimientos pertinentes a utilizar en este estudio.

Para dicho estudio webmétrico se definen dos **poblaciones**. La primera está conformada por el conjunto de los sitios web que representan a las Universidades de Latinoamérica según el Ranking Web de Universidades del Mundo de enero de 2010 (<http://www.webometrics.info/>) que cuenta con 100 sitios web, y la segunda, compuesta por el conjunto de los sitios web presentes en la Web de la UCI del año 2010 que posee un aproximado de 219 sitios web.

Esta selección de poblaciones se realiza con el fin de identificar:

1. En el caso de la primera población, el aporte positivo que brinda la utilización de los indicadores webmétricos asociados a la calidad de los sitios web representativos de estas Universidades.

2. En el caso de la segunda, la presencia o ausencia de los indicadores webmétricos indispensables para obtener una adecuada calidad del sitio web, que permita sentar las bases para el establecimiento de una estrategia metodológica que ayude a mejorar gradualmente dicha calidad en la UCI.

Estas razones dan la medida de establecer una correcta selección de las muestras y verificar que estas representen a cada una de las poblaciones, es decir, sean significativas para los intereses del estudio. Se utiliza entonces, un tipo de muestreo no probabilístico, específicamente el muestreo por cuotas, pues se tiene conocimiento fiable y seguro acerca de cuáles pueden ser los sitios web más "representativos" o "adecuados" para los fines de la investigación.

Para la observación se utilizan varias tecnologías y herramientas informáticas que posibilitan la realización del estudio webmétrico, entre las que se encuentran: Bizinformation, Web Link Validator, Emezeta Rank, Alexa y WIRE Crawler. Además se hace uso de los sistemas operativos Ubuntu GNU/Linux y Windows XP, versión 2002.

En la elaboración de la estrategia metodológica se tiene en cuenta las siguientes fases: el diagnóstico del estado actual del objeto de estudio y la intervención a partir de un conjunto de acciones en función de mejorar la calidad de los sitios web en la UCI. Las otras fases: evaluación y validación de la misma, se recomiendan para estudios posteriores de la temática, teniendo en cuenta que lo que aquí se pretende es, conocer el estado actual de los sitios web de la UCI en cuanto al comportamiento de los indicadores webmétricos, para establecer un sistema de acciones en función de mejorar la calidad de estos sitios web en la unidad de observación antes mencionada.

El contenido de la investigación se divide en tres capítulos, quedando de la siguiente manera:

**Primer Capítulo:** "Indicadores webmétricos. Aspectos teóricos generales": consta de conceptos generales y básicos que permiten comprender temas relacionados con los indicadores webmétricos, caracterización de los métodos y procedimientos utilizados por los buscadores web en el establecimiento del posicionamiento, como un indicador que influye en la calidad de los sitios web. Además se precisa el estado del arte en el ámbito nacional e internacional, así como el establecimiento de las diferentes herramientas informáticas que pueden ser empleadas para estudios webmétricos.

**Segundo Capítulo:** “Selección de los indicadores y herramientas webmétricas para evaluar la calidad de los sitios web”. Establecimiento y justificación de los indicadores webmétricos y otras herramientas informáticas seleccionados para la evaluación de los sitios web.

**Tercer Capítulo:** “Estrategia metodológica para mejorar la calidad de los sitios web en la UCI.” Fundamentación de la estrategia, valoración cualitativa y cuantitativa de los estudios realizados a los sitios web de las Universidades Latinoamericanas y la UCI. Establecimiento de un conjunto de acciones para mejorar la calidad de los sitios web en la UCI.

### **Aporte de la investigación.**

Una vez concluido el trabajo se contará con los siguientes resultados concretos:

1. Estrategia metodológica sobre la base de los indicadores webmétricos, que aporta un conjunto de acciones para mejorar la calidad de los sitios web en la UCI.
2. Artículo científico sobre la influencia o peso que tienen los indicadores webmétricos en la calidad de los sitios web.

### **Capítulo 1. Indicadores webmétricos. Aspectos teóricos generales.**

La Ciencia de la Información es un área del conocimiento que está siendo explotada por una amplia audiencia en el mundo de Internet, como resultado del continuo desarrollo humano y tecnológico. Con el surgimiento de nuevas áreas de la investigación se hace más complejo organizar y suministrar la información que los usuarios necesitan. Es por ello que, el uso de los indicadores métricos que apoyen a la toma de decisiones en variados aspectos, contribuye a evaluar las distintas fuentes de información, proporcionando una visión de la calidad actual, presentada durante el proceso de búsqueda de dicha información.

Actualmente, en la Era Digital, el impacto de la calidad de una publicación en la Web es muy importante, pues de ella depende en gran medida, que dicha publicación sea utilizada o no. Un sitio o aplicación web de baja calidad, sería perjudicial para los objetivos de la institución que lo necesite, y los usuarios insatisfechos buscarían una Web diferente que cumpla sus expectativas. Sucede que muchas personas no saben cómo balancear correctamente los indicadores webmétricos en un determinado sitio web, debido a que no tienen amplio conocimiento de los mismos. Partiendo de este análisis, es de vital importancia conocer los términos relacionados a la Webmetría y aquellos criterios formales para evaluar una publicación.

El presente capítulo contiene las definiciones y conceptos más relevantes relacionados a las métricas que permiten evaluar la calidad de los sitios web, las cuales contribuyen a la comprensión del lector acerca del tema. Se expone además, una visión general del estado del arte, así como las características generales de varias herramientas informáticas existentes que posibiliten realizar estudios webmétricos.

#### **1.1. Cibermetría y Webmetría. Surgimiento, concepto y evolución.**

El origen de la Cibermetría y la Webmetría se sitúa hacia finales de los años noventa, como resultado de los diferentes estudios métricos que comienzan a llevarse a cabo en esta época, relacionados a la naturaleza y las propiedades de Internet y de la World Wide Web (WWW) en particular. Estos estudios surgen producto al crecimiento progresivo de los recursos y fuentes de información disponibles en la Red, y por consiguiente a la necesidad de poder cuantificarlos, organizarlos e interpretarlos.

*En sus inicios fueron propuestos varios términos para designar a las nuevas disciplinas, adoptando finalmente dos, que sin llegar a serlo, a menudo se emplean como sinónimos: Cybermetrics y Webometrics. Para su traducción al español ambos fueron adaptados literalmente del inglés, dando así lugar a las expresiones Cibermetría y Webmetría. [2]*

Sería bueno resaltar que, la Cibermetría mantiene conexión con otras ciencias afines, tales como: Informetría, Bibliometría y Cienciometría. Relación entre estas, que es de solapamiento en algunas áreas, debido a que comparten un mismo enfoque y un mismo ámbito: los estudios métricos de la información. Ciencia además, que sostiene una amplia relación con la Webmetría, siendo el servicio que más se destaca en esta.

Durante la búsqueda de conocimiento referente a estos dos vocablos, variados han sido los conceptos encontrados por autores reconocidos para referirse al término de Cibermetría, entre los que se pueden citar:

Según Ali Ashgar Shiri, la Cibermetría es *“la medición, estudio y análisis de toda clase de información y medios de información que existen en el Ciberespacio y que emplean técnicas bibliométricas, cienciométricas e infométricas.” [3]*

Para Dahal *“la Cybermetrics nace con el objetivo de realizar investigaciones cuantitativas en la información electrónica del Ciberespacio.” [4]*

Isidro F. Aguillo, en su opinión aborda que la Cibermetría es *“sinónimo de Ciberinformetría y Cibercienciometría, en su sentido más genérico de disciplina, mide distintos aspectos de Internet utilizando técnicas cuantitativas bibliométricas.” [4]*

Una vez analizados los significados planteados por estos autores, se está en total acuerdo con los empleados por Shiri y Aguillo, ya que engloban el término vistos desde diferentes matices. Sin embargo, se discrepa un poco del concepto abordado por Dahal, además de que el mismo es poco abarcador. Esto se considera así, ya que analizado desde un punto de vista más general de la presente investigación, esta disciplina no nace con el objetivo de realizar investigaciones cuantitativas, sino que nace a partir de las diferentes investigaciones que comienzan a surgir sobre el Ciberespacio en los años 90, y se manifiesta entonces la necesidad de cuantificar todos los aspectos y recursos de la información desde perspectiva bibliométricas e informétricas.

Para conocer la razón de ser de la Webmetría, a continuación se muestran variados conceptos empleados por autores reconocidos para referirse a este término:

Según Abraham, la Webmetría *“estudia la relación entre los elementos principales de la Web (comparable a una red neuronal conformada por nodos y conexiones: los nodos se corresponden con los dominios, servidores y páginas que componen la Web y las conexiones con los enlaces que se establecen entre los mismos dando lugar a una matriz de conexiones que abarca todo el entorno WWW.)”* [5]

También en 1997 Tomas C. Almind y Meter Ingwerse sostienen que la Webmetría *“utiliza informaciones y datos extraídos de motores de búsqueda, tales como el número y los tipos de hipervínculos, el número y el tipo de archivos, así como la tipología del sitio web”*. [6]

Al analizar estas dos definiciones, se está de acuerdo con ambas, aunque sería bueno agregar que esta disciplina estudia muchos otros aspectos tales como: los recursos de información, estructura y tecnologías presentes en la Web, desde perspectivas bibliométricas e informétricas.

Se puede sostener entonces dos vertientes de análisis para estas dos ciencias (Cibernetría y Webmetría). La primera, con todas sus variantes terminológicas estudia la aplicación de las tradicionales técnicas informétricas a cualquier tipo de información disponible en Internet; y la segunda más restringida y que forma parte de la anterior, está basada en la aplicación de la Informetría u otras técnicas de medida a la información disponible en el entorno Web.

### **1.1.1. Aplicaciones de la Cibernetría y la Webmetría. [5]**

Los estudios relacionados con la Cibernetría se han aplicado, fundamentalmente, sobre varias áreas de la información y la comunicación. Dentro de estas se encuentran, por solo mencionar algunas: las redes de información, el correo electrónico, los recursos electrónicos y la WWW.

Esta última en particular, constituye el servicio que más popularidad le ha dado al uso de Internet, por ser el medio de soporte de los sitios o páginas web que se publican gradualmente en el Ciberespacio. Su aplicación basada en estudios cibernéticos abarca los temas relacionados con:

- ✓ El número de sitios y páginas web que existen en el mundo y su distribución por países o dominios.
- ✓ La clasificación de páginas web por tipología y lengua de los documentos.

- ✓ Las estadísticas de la utilización y de los usuarios de las páginas en un determinado período temporal.
- ✓ El número de citas que recibe cada página web.
- ✓ La posición de los sitios y páginas web más citados atendiendo a su tipología documental.
- ✓ Los tipos de recursos electrónicos disponibles en cada sitio web.
- ✓ El factor de impacto web y la productividad de los autores en la Web.
- ✓ El análisis del contenido de las páginas web.
- ✓ La identificación de publicaciones electrónicas siguiendo una distribución idiomática, geográfica o por tipologías.

Desde un punto de vista exclusivamente webmétrico, se considera que las técnicas de medición pueden aplicarse fundamentalmente en la WWW para determinar:

- ✓ Propiedades gráficas de la Web: Se miden las propiedades gráficas a escala macroscópica y microscópica mediante las relaciones entre las páginas web y sus enlaces.
- ✓ Importancia de las páginas web: Esta métrica formaliza las nociones de calidad y relevancia de las páginas web teniendo en cuenta las demandas de los usuarios.
- ✓ Caracterización del uso: Mide los patrones de búsqueda que utilizan los usuarios para localizar fuentes web.
- ✓ Similitud de páginas web: Cuantifica el grado de relación entre las páginas web teniendo en cuenta indicadores como el análisis de citas.
- ✓ Búsqueda y recuperación de páginas web: Considera las técnicas para evaluar y comparar el rendimiento de los servicios de búsqueda y recuperación de información web.
- ✓ Teórica de la información: Mide los aspectos relacionados con las necesidades, producción y consumo de la información web.

A pesar de las aparentes diferencias entre las áreas cuantificables desde el punto de vista de la Cibermetría y la Webmetría, todas las definiciones coinciden en un aspecto clave: los estudios cibernéricos y webmétricos tratan de “medir” la información disponible en Internet y la WWW respectivamente. Para esto se basan en la utilización de los llamados indicadores cibernéricos.



## 1.2. Indicadores cibernéticos.

Los indicadores son elementos básicos que brindan información cualitativa y/o cuantitativa del control de cómo funciona una actividad, pues hacen referencia a parámetros estables que sirven de magnitud de comprobación del funcionamiento de esta. Los indicadores deben ser: accesibles (fáciles de identificar y recopilar), pertinentes (para lo que deseamos medir), fieles (que informen con fidelidad de las condiciones de los datos que se recogen), objetivos (no ambiguos en su interpretación), precisos (para la acción que se quiere estimar), unívocos (parámetros exclusivos de lo que se mide) y sensibles (que permitan recoger y estimar variaciones de aquello a lo que referencian).

Los indicadores webmétricos más conocidos pueden ubicarse dentro de las siguientes categorías:

**Indicadores descriptivos:** Miden fundamentalmente el tamaño o número de objetos encontrados en cada sede web dígase: riqueza de páginas, ficheros media o ricos, densidad de enlaces totales y únicos.

**Indicadores de impacto y visibilidad:** Miden el número total de citas o enlaces que recibe el sitio y el número de enlaces que apunta a una sede web respectivamente. El número y diversidad de enlaces externos recibidos, el volumen de los mismos respecto a los contenidos objeto de enlace (llamado factor de impacto web) o índices que se construyen de acuerdo al peso relativo de las sedes de origen de los enlaces, permiten establecer listados ordenados de acuerdo a la jerarquía numérica de estos indicadores.

**Indicadores de popularidad:** Miden en términos de números y características las visitas recibidas a una sede web.

La gran variedad de estos indicadores, posibilita el desarrollo de estudios comparativos que abarquen campos como: dominios geográficos y funcionales, instituciones públicas y privadas, documentos electrónicos, espacios web científicos y comerciales, contextos sociales, regiones geográficas, etc. Esto hace difícil definir sus características básicas.

A continuación se realiza una agrupación de los mismos según las variables o categorías más generales que permiten evaluar.

Categoría	Indicador
-----------	-----------

Infraestructura	Número de host, de servidores web, de usuarios, de dominios, de sitios, de sitios institucionales, etcétera.
Tamaño	Número de páginas, de objetos, de objetos multimedia, de archivos ejecutables, tamaño de los archivos, distribución por lenguajes, evolución temporal, número de niveles, de enlaces por página, etcétera.
Calidad	Porcentaje de enlaces válidos, de errores de enlace, apariencia, etcétera.
Conectividad	Total de enlaces, de enlaces por página, número de enlaces internos, de enlaces externos, etcétera.
Visibilidad	Número de enlaces recibidos o externos, enlaces nacionales externos, enlaces internacionales externos, etcétera.
Impacto	Factor de impacto.
Popularidad	Número de visitas.

**Tabla 1: Indicadores por categoría.**

### Utilidad de los indicadores: [7]

- ✓ **Elemento de planificación:** Durante los procesos de planificación se utilizan con frecuencia los indicadores para establecer la meta u horizonte a donde se quiere llegar. También son usados como estándar o medio de verificación o evaluación de un proceso.
- ✓ **Estándar de seguimiento y control:** El indicador ayuda a entender o muestra el estado del problema, sirve de termómetro de una problemática o fenómeno, permite identificar qué tan lejos se está de donde se desea estar, es decir, ayuda a determinar la brecha entre lo planificado o esperado y el punto actual en el que se hace la valoración o medición. Un buen indicador es una señal de alerta antes que el problema sea irremediable e induce a reconocer que es necesario resolver dicho problema.
- ✓ **Herramienta para la toma de decisiones:** Son factores para establecer el logro y el cumplimiento de la misión, objetivo y metas en determinado proceso, proporcionando la información de apoyo para la toma de decisiones y el planteamiento de políticas y estrategias para minimizar el problema.

## Ventajas Prácticas.

Las ventajas de los indicadores web están dadas por los siguientes aspectos.

- Mayor masa crítica para análisis de patrones globales y sectoriales:
  - ✓ Mejor tratamiento con técnicas estadísticas.
  - ✓ Nuevas unidades: Mayor finura en el análisis.
  - ✓ Perspectiva cuantitativa y objetiva.
- Mejores resultados esperados:
  - ✓ Batería de indicadores más amplia.
    - Medidas combinadas.
    - Visualización más espectacular.
  - ✓ Seguimiento individualizado.
  - ✓ Medidas directas e indirectas.
  - ✓ Comparación con descriptores “offline”.
- Ventajas políticas:
  - ✓ Medida de la producción de conocimientos.
    - Incremento de los contenidos.
  - ✓ Mayor visibilidad social.

### 1.3. Conceptos relacionados con la calidad.

La calidad es un término que tiene múltiples definiciones por la evolución que ha tenido y su planteamiento por diferentes autores. Su significado representa más bien, un conjunto de propiedades y características específicas de un objeto, que tienen la capacidad de satisfacer necesidades y lograr conformidad por parte de los consumidores. También se encuentra estrechamente relacionada con la mejora continua en cualquier organización y que afecta a todas las personas y a todos los procesos; por lo cual se deduce que la calidad es responsabilidad de todos.

Según algunos expertos del mundo, la calidad se define:

- Philip Crosby: *“Calidad es cumplimiento de requisitos.”*[8]

- Joseph Juran: *“Calidad es adecuación al uso del cliente.”*[8]
- Armand V. Feigenbaum: *“Satisfacción de las expectativas del cliente.”* [8]
- Genichi Taguchi: *“Calidad es la menor pérdida posible para la sociedad.”* [8]
- Walter A. Shewhart: *“La calidad como resultado de la interacción de dos dimensiones: dimensión subjetiva (lo que el cliente quiere) y dimensión objetiva (lo que se ofrece).* [8]

## **Calidad de la información.**

Toda institución para su desarrollo, necesita información que le posibilite tomar decisiones que dependen de la validez de la información utilizada. El valor y la credibilidad del contexto de la información son los aspectos más importantes a tener en cuenta para divulgar algún contenido en la Web, pues del contenido presentado también depende el triunfo o fracaso de un sitio web.

Un factor significativo para el éxito de los sistemas informáticos lo constituyen la calidad de los datos procesados; esto infiere a que los datos no contengan errores, sean eficientes y estén actualizados. También deben cumplir ciertas normas mínimas de calidad entre las que se destacan las siguientes: los datos deben ser únicos, deber ser veraces y ser flexibles.

La calidad de la información depende también del tipo de Web donde esté contenida. Si la Web es invisible, entonces la información contenida en ella será invisible también porque el acceso a esta información está restringido para los usuarios que no conozcan de la presencia de la misma.

## **Calidad de los sitios web.**

Antes de analizar el término relacionado a la calidad de los sitios web, se hace necesario comprender: **¿Qué es un sitio web?** Para ello se citarán, algunas definiciones por autores reconocidos:

*Aguillo: Lugar físico interconectado a la Red de Internet donde se encuentra almacenada la información en formato electrónico accesible mediante el protocolo HTTP, es decir, un ordenador que actúa como servidor web con una dirección IP propia.* [5]

Koehler: *Colecciones de una o más páginas web que comparten algún tema o principio organizativo en común y que se conectan a través de enlaces hipertextuales. Aunque a menudo incluye un dominio completo de nivel servidor (SLD), los servidores pueden alojar más de un sitio web con el mismo SLD. [5]*

A través del análisis de estas definiciones, se puede concretar de que un sitio web no es más que: un conjunto de páginas interrelacionadas mediante enlaces hipertextuales, que se muestran a través de Internet con varios propósitos: presentar información sobre un tema, instruir sobre un tema, hacer publicidad, distribuir materiales, etcétera.

Para lograr estos propósitos se necesita de un sitio que contenga la calidad requerida, es decir, un conjunto de características que haga al sitio más visible, navegable y consistente. Algunas de ellas son: contenido entendible y no muy extenso, presentación con apariencia profesional, información valiosa en la toma de decisiones, menús de acceso suficientes para guiar a los usuarios, entre otros. De tal forma que la apreciación del visitante sobre el sitio, sea de un nivel significativo.

Los enlaces que conectan los documentos web, también son de gran importancia para la calidad de los sitios web. La conectividad de enlaces entre las páginas web, contiene gran cantidad de información implícita de unos enlaces con respecto a otros. Así los documentos de mayor calidad, que contienen información clara, precisa y útil, son enlazados por muchas otras páginas. De esta forma un sitio o página web obtiene más popularidad por la calidad de sus enlaces y la cantidad de veces que aparece referenciado en otros sitios web, así como en buscadores.

Según Chu, *cuando se realiza un estudio de la calidad de los enlaces hipertextuales de las colecciones digitales, es preciso tener en cuenta exclusivamente los enlaces basados en el contenido. [5]*

## **1.4. Motores de búsquedas o buscadores. Su funcionamiento.**

El estudio del posicionamiento de páginas o sitios web en los motores de búsqueda es una de las técnicas más novedosas de la Cibermetría, pues se ha convertido en el principal intermediario de la recuperación de información en Internet.

La Web constituye la vía más rápida para buscar información sobre algún tema específico. Los usuarios generalmente se apoyan en la utilización de motores de búsqueda, pues cuando una persona necesita de información sobre un tema específico, y no conoce o no recuerda la dirección de una página web, siempre

recurre a la ayuda de los buscadores. Estos no son más que páginas web donde en realidad se consulta una base de datos, haciendo más rápido y fácil la obtención de la información requerida.

Existen diferentes tipos de motores de búsqueda o buscadores, desde los más hasta los menos usados: Google, Yahoo, Microsoft, MSN/Windows Live, AOL Media Network, Fox Interactive Media, YouTube, eBay, Wikipedia y Amazon, entre otros.

Dentro de los dos grandes grupos de buscadores díganse nacionales e internacionales, es posible conseguir el lugar que cada sitio web se merece. Esto depende del trabajo y voluntad que cada administrador web ponga sobre el sitio para alcanzar el triunfo. Además, se necesita conocer acerca del funcionamiento de los buscadores y los algoritmos que emplean para posicionar los sitios web.

#### **1.4.1. Funcionamiento de los motores de búsquedas. Relación con los sitios web.**

El proceso de búsqueda en los buscadores comienza a partir de la entrada de una palabra o frase, las cuales son palabras claves utilizadas por el buscador como referencia para realizar la búsqueda. En respuesta se muestra un listado de páginas web ordenada por los criterios del buscador. Generalmente los primeros 10 o 15 resultados de la lista son los de mayor interés para el usuario, pues es muy molesto para este seguir la búsqueda en páginas que se encuentran en bajas posiciones. De esta forma cada sitio web compite con miles de miles que tienen las mismas palabras claves, y solo tienen posibilidades de ser visitados los primeros de la lista de resultados mostrada.

Un sitio web aparece o no aparece entre los primeros lugares del resultado de una búsqueda, en dependencia del software con el que fue implementado el buscador utilizado. Este programa de búsqueda obedece a una serie de criterios y funciones organizadas en un algoritmo particular, el cual se encarga de calcular el lugar más adecuado en el que debe mostrarse una página web. *Los algoritmos de búsqueda permiten responder a una consulta, con una lista de sitios web cuyo contenido es verdaderamente relevante para quien hace la consulta. La tecnología del buscador conjuga diversos criterios cualitativos en el contenido de los sitios web existentes, con los que verifica si son pertinentes o no para el usuario. [9]*

Cada buscador tiene su propio algoritmo de búsqueda, es por eso que no todos operan de la misma forma. Solo algunos de los motores de búsqueda valoran meta-tags, títulos, textos alternativos a las imágenes, popularidad de los enlaces, frecuencia de click-through, número de veces que las palabras

clave aparecen en el texto, etcétera.

Por ejemplo, Google utiliza un robot llamado Googlebot que se encarga de rellenar en una base de datos primaria el código de las páginas que visita. Luego un motor de búsqueda se encarga de consultar esa base de datos una vez realizada la pregunta, y posteriormente muestra los resultados dado un complicado algoritmo, que tiene la función de clasificar una página por temas, origen, etcétera, listándolas ordenadamente de acuerdo con su mayor o menor relación con el término buscado por el usuario.

Cuando un robot visita una página web, busca identificar sobre el código de dicha página una serie de zonas: la cabecera, el cuerpo, las etiquetas clave, entre otros. Estos criterios que utiliza el robot sobre las diferentes partes del documento HTML<sup>1</sup> de la página son aspectos muy importantes a tener en cuenta a la hora de posicionar un sitio web, ya que estos constituyen las palabras claves que ayudan a que el sitio web sea encontrado. En conjunto, se realiza una serie de análisis sobre las áreas del documento verificando si estas cumplen con determinadas variables como son: frecuencia, relevancia, peso y recuento.

Por todo lo explicado anteriormente, es necesario conocer qué valora cada buscador y tener las herramientas necesarias para optimizar la programación y el diseño de un sitio web completo para su indexación en los buscadores. Así, el sitio web será diseñado y desarrollado de manera tal que le facilite el trabajo a los buscadores y a la vez a los usuarios, ya que en la medida que la página web responda a los criterios con los que funcionan los motores de búsqueda, tendrá más posibilidades de aparecer entre los primeros puestos del resultado de una búsqueda.

## **1.5. Herramientas Webmétricas.**

La combinación de diferentes herramientas, necesarias para la realización de los estudios webmétricos, permite obtener datos cuantitativos de los sitios web previamente identificados. De esta forma los datos obtenidos son más fiables para alcanzar resultados concretos. Para lograr un estudio webmétrico satisfactorio es imprescindible estudiar las distintas herramientas posibles a emplear, que logran además, facilitar el trabajo.

---

<sup>1</sup> HyperText Markup Language. Lenguaje de marcado predominante para la elaboración de páginas web.

## Verificadores de enlaces.

- **Web Link Validator.** [10] [11]

El Validador del acoplamiento del Web es una poderosa y completa herramienta de administración de sitios y verificación de enlaces, que ayuda a los administradores de sitios a automatizar el proceso de prueba de sus sitios y de los sitios de Internet. Este es además, un software para escritorio que verifica la integridad del sitio, valida la sintaxis, creación automática de informes y funciona sobre Windows (98, ME, NT, 2000, XP, 2003, Vista).

El programa realiza un análisis profundo de las páginas de los sitios de Internet, con el propósito de encontrar enlaces rotos, incluso si se utiliza Java Script o Flash en dichas páginas. A pesar de lo que sugiere el nombre de esta herramienta, el Web Link Validator no solo se limita al control de los vínculos, sino también al problema de los archivos huérfanos, errores en el código HTML, carga lenta obsoleta y minúsculas páginas.

El software ayuda a mantener el contenido de alfabetización en el lenguaje humano, utilizando el built-in corrector ortográfico encuentra y corrige cualquier error de ortografía en inglés y en otros 20 idiomas. El programa además, puede tener acceso a recursos de HTTPS y a las páginas protegidas con contraseñas, autenticación basada en la Web y los servidores proxy si está requerido.

Después de realizar el análisis en profundidad del sitio, Web Link Validator genera un informe sobre las páginas revisadas, siendo mostrado en un formato HTML, TXT, RTF, CSV o MS Excel perfectamente legible. Luego, los poderosos filtros permiten una investigación profundizada de cualquier problema que haya sido descubierto, lo cual ayuda a los administradores de sitios concentrarse en los aspectos específicos del problema para poder resolverlos.

Otras funciones de Web Link Validator incluyen la habilidad de verificar la copia de un sitio de Internet guardada en un disco duro local, configuraciones individuales para cada sitio, envío automático por correo electrónico de los reportes del sitio, un motor de verificación alternativo basado en Internet Explorer y una interfaz de programa personalizable.

Esta poderosa herramienta es una solución clave para los administradores de sitios, desarrolladores de Internet, y cualquier otra persona que desee mantener la calidad de los sitios



web. Web Link Validator mejora la experiencia del visitante y mejora su confianza con un sitio web, ayudando a los propietarios del sitio ganarse la confianza y la lealtad de sus clientes.

- **Content Analyzer.** [12]

Content Analyzer es un analizador profesional de contenidos de los sitios web. Esta herramienta especializada para el análisis, examina páginas HTML en línea o fuera de línea, con el propósito de brindar un informe detallado acerca de la densidad de las palabras clave, su peso y relevancia. Incluso puede sugerir palabras claves o frases que se pueden usar para atraer visitantes más calificados a un sitio.

En las ayudas del Content Analyzer, se pueden encontrar las palabras claves más populares de la página web y las palabras débiles y sin optimización. La herramienta trabaja con palabras solitarias y también con palabras claves, descubre que páginas están sub-optimadas, investiga los links dentro del propio sitio web, analiza otros sitios web para aprender de estos sus palabras claves, pues tiene la capacidad de analizar sitios web alejados y mucho más.

El uso de esta aplicación de software ayuda a obtener los puestos más altos para un sitio web en los buscadores. Con la herramienta Content Analyzer se puede además, obtener un informe en tiempo real acerca de las posiciones actuales de un sitio web y también las de sus competidores para cualquier palabra clave o frase de búsqueda. Se puede alterar el contenido o las palabras elegidas de un sitio para mejorar su posición antes de que los competidores realicen otra movida.

Content Analyzer usa un algoritmo inteligente que toma en cuenta la localización de etiquetas y palabras claves. Esta herramienta puede analizar el sitio web entero o las páginas seleccionadas, generando informes fáciles de comprender y estructurados en un árbol.

El programa soporta Unicode y puede usarse para analizar los contenidos de una página web en cualquier lenguaje.

- **W3C Link Checker.** [13]

El W3C Link Checker es un validador de enlaces desarrollada por W3C, mediante el cual se puede comprobar si existen enlaces, anclas (hipervínculos) y objetos de referencias rotos en un sitio o páginas web.

Esta herramienta posibilita obtener información profunda sobre los vínculos en una página web, acerca de los errores en las direcciones URL, la inexistencia de enlaces y otras cuestiones como los problemas del lado del servidor. Tiene varias opciones para el control de páginas individuales o de control recursivo a través de sitios web enteros y maneja redirecciones. Presenta grandes beneficios a quienes lo usan, por ser una herramienta libre, además de que es rápida y fiable.

Para su uso se le indica la dirección URL de la página que se quiere validar, y el sistema se conecta a todos los enlaces mostrando el tiempo que tarda en recuperar cada página, y muestra al final un resumen con los enlaces rotos, indicando la línea donde está el enlace, el código de error obtenido, y la descripción del error devuelta por el servidor. También se le puede definir un grado de recursión, para que compruebe la validez de los enlaces hasta un determinado nivel de profundidad.

El programa es parte de los validadores del W3C y herramientas de calidad web, que ayuda a mejorar la calidad de un sitio completo.

### **Volcadores de sitios web.**

*Los Volcadores de sitios web, son programas que permiten el volcado automático selectivo o general de páginas o sedes web. A continuación se dan a conocer algunas de estas herramientas. [1]*

- **Website Extractor 9.07. [14]**

Es un software diseñado para descargar sitios web completos en un ordenador. Contiene un navegador integrado, el cual incluye una doble pestaña que le permite explorar todos los sitios web tanto online como offline (en su propio disco duro).

Mediante este programa se puede extraer todos los elementos que componen un sitio web para ser visualizado sin conexión a Internet, en una interfaz. Esta interfaz muestra un mapa del sitio, que ayudará a navegar entre los contenidos descargados pudiendo ser visualizados en el navegador interno antes mencionado. Todos los datos obtenidos se almacenan en un directorio seleccionado y contienen únicamente los archivos y directorios que coinciden con las instrucciones del filtro que se le indicó.

El Extractor es esencialmente un robot de búsqueda y fue diseñado para una navegación con

seguimiento rápido a través de los enlaces del Ciberespacio, descargando páginas web bajo pedido del usuario. Ofrece numerosas configuraciones y opciones para facilitar esta tarea.

Una particularidad del Website Extractor es que permite configurar las descargas al gusto y a las necesidades del usuario, filtrando los ficheros por extensión o tamaño (mínimo o máximo), indicando hasta qué nivel de la Web se quiere descargar, etc. También permite descargar de forma automática cualquier archivo que no pudo ser copiado, debido a errores en la transferencia o malas conexiones.

El programa se encuentra equipado para correr a través de un servidor proxy y descargar solo archivos revisados o nuevos, ignorando documentos que ya han sido copiados.

- **Teleport Pro. [15]**

Teleport Pro es una herramienta multipropósito de alta velocidad, que busca y captura ficheros de Internet, gráficos y puede además realizar descargas de sitios web completos. Despliega hasta 10 hilos de recuperación simultánea, accede a sitios protegidos por contraseña, filtra archivos por tipo y tamaño, busca contraseñas y mucho más.

Para realizar la búsqueda se introduce la dirección URL de la página, y la aplicación se encarga de realizar un escaneo determinando todos aquellos ficheros que sean descargables. Los ficheros capturados se colocan en una estructura tipo directorio para navegar o para su uso individual. Posee un periodo de evaluación de 50 ejecuciones, pudiendo descargar solamente 500 ficheros por ejecución. Dentro de las características más conocidas de Teleport Pro se encuentran:

- ✓ Descargar todo o parte de un sitio web en el ordenador.
- ✓ Crear un duplicado exacto de un sitio web, incluyendo la estructura de directorios.
- ✓ Buscar archivos de un tamaño o tipo específico en un sitio web.
- ✓ Descargar un listado de archivos en una dirección conocida.
- ✓ Explorar cada sitio web que está enlazado desde un sitio web central.
- ✓ Buscar palabras clave en un sitio web.

- **Webcopier. [16]**

Es una herramienta poderosa que crea copias de sitios web y las almacena localmente en un

sistema, para luego navegarlos offline. WebCopier está compuesto por una gran cantidad de funciones que te permite realizar la descarga de un sitio web de manera totalmente personalizada, incluso si solamente deseas descargar archivos específicos como imágenes, fotografías y demás.

Entre sus características principales se destacan las siguientes funcionalidades: permite descargar hasta 100 archivos simultáneamente, puede imprimir todo el sitio o una parte específica de ello, soporta servidores proxy, permite especificar un nombre de usuario y contraseña para acceder a sitios web seguros, posee una opción para bloquear todas las ventanas emergentes o pop-up publicitarios; almacena todos los caminos en forma relativa, de modo que los ficheros se pueden mover fácilmente a otro ordenador, CD, dispositivo de seguridad.

### **1.6. Estudios webmétricos internacionales y nacionales.**

#### **España. [17]**

SEOLUCION, la división de marketing en buscadores de Adesis Netifle (Compañía pionera en el desarrollo de Internet de tercera generación) elaboró su primer estudio de posicionamiento optimizado en buscadores en el sector de hoteles, en el que se destacaba la escasa popularidad online de los principales establecimientos hoteleros de este país. A partir del estudio de competidores de esta división, se reflejó el nivel de visibilidad de las páginas web de las principales cadenas hoteleras de España, determinado por la posición de las mismas en los buscadores de Internet. El estudio, que además abarcó el análisis de otros indicadores web como fueron: popularidad, saturación, ranking de página y ranking Alexa, determinó que hoteles principales como el HUSA, tuvieran el portal más visible del sector con un índice de visibilidad de un 45% y NH Hoteles presentara un índice de visibilidad del 35% y se destacara por ser el portal con más popularidad.

#### **Brasil. [18]**

El estudio consiste en la investigación del campo de la Webmetría, definiendo el lugar que esta ocupa dentro de la Ciencia de la Información, así como el análisis de los flujos de información que se dan en la Web, mediante la aplicación de diferentes indicadores webmétricos a los sitios de los centros afiliados al Consejo Latinoamericano de Ciencias Sociales (CLACSO). Para la extracción, cuantificación y representación gráfica de los datos se utilizaron los motores de búsqueda (Google, Yahoo y Alta Vista), programas mapeadores (Xenu Link Steuth), y programas de análisis y visualización de redes (Ucinet 5 y

Netdraw).

Los resultados obtenidos:

1. Indicaron que la Informetría debe ser entendida como el campo más amplio, que incluye además los campos de la Cibermetría y la Webmetría.
2. Presenta la posición de que la Webmetría no está totalmente absorbida por la Bibliometría, dado que la información en el ambiente digital no puede ser considerada como registrada, sino más bien disponible en soporte electrónico.
3. La aplicación de los indicadores webmétricos demostró que el Factor Impacto Web y Densidad por Link no ofrecen resultados satisfactorios. Los demás indicadores demuestran su utilidad y relevancia para la medición de los datos recolectados en la Web.

### **Cuba. [19]**

El estudio de la visibilidad de los centros del MES de Cuba, a partir de un análisis webmétrico, fue realizado en los 17 centros que tienen sitios web y se determinó la relación entre el número de enlaces y el tráfico de las páginas de los centros del MES a partir de los análisis de los buscadores Google, AlltheWeb y Altavista, de Internet Archive y el Ranking Alexa. El comportamiento de estas páginas web fue analizado en Internet Archive desde enero de 1996 hasta febrero del 2007.

Los resultados obtenidos fueron:

1. En la comparación entre los buscadores (Google, Alltheweb, Altavista), la mayor media de los enlaces de los centros del MES (732,53) se obtuvo del buscador Altavista con  $p < 0.01$ .
2. En la medida que se incrementaron las páginas archivadas en Internet Archive al paso de los años, mayores citas (o enlaces) recibieron los sitios web de los centros del MES, de los buscadores, en especial de Google.
3. Los enlaces y el ranking de los centros del MES se comportaron significativamente mejor en el buscador Google para un EE de  $\pm 0,01$  y un R2 igual a 0,12 %.
4. Existen diferencias significativas en las páginas de los centros del MES en Internet Archive respecto a los enlaces en el Ranking Alexa.

A través del estudio realizado, se ha presentado en este capítulo una amplia gama de información que ayuda a la realización de la tesis. Los conceptos fundamentales relacionados con el objeto de estudio de

este trabajo, contribuye a la correcta comprensión del mismo, logrando de esta forma una adecuada selección de los indicadores webmétricos a utilizar en la evaluación de la calidad de los sitios web. La exhaustiva investigación llevada a cabo, acerca de las posibles herramientas cibernéticas a utilizar, divididas en dos categorías: verificadores de enlaces y volcadores de sitios, sirven de gran ayuda para el análisis y selección de aquellas herramientas que brinden los elementos necesarios para la realización de este estudio.

Finalmente completados los primeros objetivos, se establecen las condiciones para continuar con el desarrollo del trabajo.

## Capítulo 2. Selección de los indicadores y herramientas webmétricas para evaluar la calidad de los sitios web.

Como se viene analizando los indicadores, por lo general, son elementos informativos del control de cómo funciona una actividad, pues sirven para indicar la existencia o ciertas características de un hecho no observable. Se utilizan para describir y evaluar los diferentes aspectos de una situación actual o comparar una situación común con relación a una situación deseada.

En el presente capítulo se dan a conocer los indicadores cibernéticos y webmétricos que permiten caracterizar los sitios web, así como un grupo de herramientas cibernéticas que posibiliten la realización de estos estudios. Se seleccionan y justifican aquellos indicadores webmétricos que se consideran relevantes para obtener una buena calidad de los sitios web, y las herramientas informáticas que permiten evaluar estos sitios web.

### 2.1. Indicadores webmétricos empleados en los análisis métricos.

Los indicadores métricos más conocidos utilizados en el análisis métrico de los recursos digitales, pueden ubicarse dentro de la siguiente lista de categorías.

- ✓ Indicadores de tipos institucionales: Cantidad de páginas en determinados sectores.
- ✓ Indicadores regionales.
- ✓ Indicadores idiomáticos.
- ✓ Indicadores de tipología de sitios: Sitios académicos, comerciales, de sectores públicos o privados.
- ✓ Indicadores de tamaño: En sus dos variantes, tamaño documental: número total de páginas comprendidas en un dominio o tamaño informático: tamaño en bytes de una sede web.
- ✓ Indicadores de densidad: También tiene dos variantes, densidad hipertextual: media de enlaces por página, y densidad multimedia: media de objetos multimedia por página.
- ✓ Indicadores de profundidad: Número máximo de niveles de una sede.
- ✓ Indicadores de luminosidad: Total de enlaces emitidos desde una sede.
- ✓ Indicadores de visibilidad: Número total de enlaces externos diferentes recibidos por una sede, existen variantes para calcular la visibilidad nacional con límite a los enlaces recibidos en el mismo país, etcétera.

- ✓ Indicador de navegabilidad: Número total de enlaces internos respecto al total de páginas.
- ✓ Indicadores de validez hipertextual: Porcentaje de enlaces válidos respecto al total.
- ✓ Indicadores de cooperación (colegios invisibles).
- ✓ Indicador de diversidad: Distribución de las características de los enlaces recibidos por una página.
- ✓ Medidas de popularidad: Número y distribución de las visitas recibidas en un plazo determinado.
- ✓ Indicadores de impacto: Resultado de dividir el número total de enlaces externos diferentes recibidos por una sede entre el tamaño de esta expresado en número de páginas.
- ✓ Indicadores para el estudio del comportamiento de usuarios en la recuperación de información.

### **2.1.1. Indicadores webmétricos más empleados para la evaluación de los sitios web.**

La evaluación de los recursos digitales es un proceso importante e imprescindible en el ambiente informativo, para el correcto funcionamiento de las diversas actividades institucionales. El objetivo fundamental de la evaluación radica en brindar información confiable y útil para conocer el desenvolvimiento de una actividad, con el propósito de elevar su nivel de calidad.

A continuación se muestra un conjunto de indicadores que se pueden emplear para la toma de decisiones sobre la calidad de productos o servicios de información, particularmente para la evaluación de los sitios web en Internet.

- Indicadores de contenido:
  - ✓ Información contrastada y rigurosa, editada o supervisada.
  - ✓ Existencia de información valiosa.
  - ✓ Volumen de información suficiente.
  - ✓ Información actualizada.
- Indicadores de navegación y recuperación:
  - ✓ Sistema de búsqueda: herramienta que permita la búsqueda de información existente en el sitio.
  - ✓ Mapa de navegación o sumario: representación gráfica de la estructura de un sitio web.
  - ✓ Recorrer la estructura del sitio web sin perderse.
  - ✓ Opciones de navegación claras.
- Indicadores de luminosidad (calidad de enlaces externos):



- ✓ Enlaces externos.
- ✓ Enlaces evaluados y seleccionados.
- ✓ Enlaces actualizados: ritmo o frecuencia de actualización de un recurso.
- ✓ Valor añadido en los enlaces.
- Indicadores de visibilidad (número de enlaces que recibe de otros recursos):
  - ✓ Enlaces desde otras páginas.
  - ✓ Título del sitio web en la barra del navegador.
- Indicadores de disponibilidad y accesibilidad:
  - ✓ Información en diversos idiomas.
  - ✓ Posibilidad del sitio de aparecer por su temática en los resultados de otros motores de búsqueda.
  - ✓ Opción de búsqueda que permita moverse con facilidad en el sitio.
  - ✓ Información gratuita.
- Indicadores de diseño y cobertura:
  - ✓ Nivel de atracción del diseño.
  - ✓ Imágenes fáciles de bajar.
  - ✓ Facilidad de la lectura (color, fondo, gráfico).
  - ✓ Gráficos y diseño sugerente.
  - ✓ Facilidad de arquitectura.
  - ✓ Equilibrio entre imagen y texto.
  - ✓ Menú principal en todas las páginas del sitio.
- Indicadores de confiabilidad:
  - ✓ Fecha de actualización del sitio.
  - ✓ Seguridad de la validez de sus enlaces (páginas inexistentes, cambios de dirección).
  - ✓ Forma de verificar la legitimidad de la organización (correo postal, teléfonos, fax, e-mail, etcétera).
  - ✓ Datos sobre resultados de la organización.
  - ✓ Código de ética de la organización.
- Indicadores de profesionalidad:
  - ✓ Lenguaje preciso y conciso.

- ✓ Refleja propósitos profesionales.
- Indicadores de cobertura y objetividad:
  - ✓ Relación entre los objetivos de la organización y los contenidos de la fuente.
  - ✓ Información debidamente citada.
  - ✓ El sitio cubre diversos aspectos sobre un tema.
  - ✓ Alcance internacional.
  - ✓ Enlaces complementarios al contenido de la fuente.
- Indicadores de valor agregado:
  - ✓ Servicios en línea.
  - ✓ Formularios para interactuar con la organización.
  - ✓ Sección de noticias diarias para los clientes.
  - ✓ Información sobre mecanismos de ventas.
  - ✓ Sección de quejas y sugerencias.

Es oportuno señalar que, estos indicadores son solo algunos seleccionados a través de la unificación de criterios de personalidades reconocidas, y que pueden servir de ayuda para la evaluación de los sitios web. Esto indica que no son los únicos, ya que el tema va más allá del alcance de solo un grupo de trabajo.

### **2.1.2. Indicadores webmétricos relevantes para estimar la buena elaboración de un sitio web.**

Cuando se quiere que un sitio web luzca excelente con cualquier visualizador, se necesitan tener en cuenta algunas reglas o criterios que ayuden a elaborar un sitio web con apariencia profesional. A continuación se explican los indicadores que se consideran más importantes para conformar un buen sitio.

#### **Indicador de consistencia:**

La consistencia da lugar a una unidad de imagen y de estructura a la información de las páginas del sitio web. La imagen se relaciona con el colorido, la iconografía, composición o disposición visual y orgánica de los elementos. La unidad de imagen se logra cuidando la homogeneidad (unidad) y equilibrio de los elementos anteriores. La estructura se forma con la organización y relación orgánica de las partes y los contenidos de información, con la jerarquización y distribución de la información.

### **Indicador de navegación:**

La estructura de la información corresponde a la organización de esta, dando lugar a un esquema de navegación o exploración que debe ser consistente y sencillo para permitir al usuario obtener la información deseada fácil y rápidamente, evitando así el exceso de páginas o de redundancia (repetición de la información).

Para superar la repetición de contenidos se emplean las ligas de hipertexto a lugares donde la información se encuentra desarrollada. Para organizar la información intervienen los menús y las barras de navegación.

### **Indicador de punto de enfoque:**

El diseño de una página debe dirigir la atención (enfocarla) a los aspectos más importantes de la página o de la información que contiene. Los GIF<sup>2</sup> animados o activos se emplean para resaltar esos aspectos importantes, sin saturar la página de ellos.

### **Indicador de colorido:**

Es necesario cuidar la compatibilidad entre la paleta de colores del visualizador y la empleada para elaborar la página. Esta compatibilidad se cuida, especialmente, cuando la configuración de los monitores varía entre 16 y 33,000 colores o más. A más imágenes, más colores se emplean, pero no la mayoría de los colores desplegados dependen de la paleta de colores del visualizador (Netscape tiene una paleta de 216 colores).

### **Indicador de tipografía:**

Es recomendable usar una o dos tipografías (en tamaño y estilo). No es buena la diversificación.

### **Indicador de diseño o formato:**

El diseño tiene que ver con el formato. El tamaño del formato debe cuidarse para que quepa en la pantalla del visualizador, ya que las barras de desplazamiento y de herramientas de este ocupan parte de la pantalla.

---

<sup>2</sup> Graphics Interchange Format. Formato de compresión de imagen limitado a 256 colores.

El formato corresponde a la organización de la información que aparece en una pantalla (generalmente la mitad de la página). La página web se concibe por pantallas, para darle unidad al formato.

Generalmente, el tamaño de una pantalla debe conservar el formato de 640x480 pixeles (o el formato de monitores más antiguos) y dentro de este tamaño de espacio, concebir el diseño de la página.

### **Indicador de imágenes:**

El empleo de imágenes tiene relación con la limpieza de la presentación, la agilidad de la página y el tiempo en cargarse. En una página web hay que cuidar el tamaño de las imágenes y la cantidad que se integra en la página para que el servidor que la soporta no se sobrecargue o saturar, ni la conexión sea demasiado lenta.

El tamaño se considera de dos maneras:

1. Por el espacio que ocupa visualmente en la página (se mide en pixeles de ancho y de alto).
2. Por el espacio que ocupa en el archivo del disco duro donde se almacena (se mide en bytes; un byte equivale a 8 unidades de bit, menor unidad de medida existente).

El espacio que ocupa una imagen depende de la información que almacena (colores y definición). De este modo, imágenes del mismo tamaño pueden ser diferentes en el espacio que ocupan en el disco. Por esta razón el tamaño ideal de las imágenes para una página web y el número de ellas no debe pasar de 50K por página (K=Kb significa kilobytes, aproximadamente 1,000 bytes).

### **Indicador de tamaño de páginas:**

El tamaño óptimo de una página es de dos o tres pantallas (una cuartilla o cuartilla y media). No debe ser mayor al equivalente de cinco cuartillas (cada cuartilla equivale a dos pantallas). Un documento grande ha de descomponerse en varias páginas, en unidades de contenido de una y hasta cinco pantallas.

Para este efecto, se puede pensar en una página web como el espacio continuo que se despliega para mostrar información y que se puede recorrer a manera de rotafolio, con el auxilio del scrolling (enrollador o barra de desplazamiento) colocado a la derecha de la página (identificado por flechas).

### **Indicador de estructura o ramificaciones:**

Esta característica tiene que ver con la organización de contenidos por unidades de información. Es decir, por temas o módulos de información y los niveles en que se coloca esta.

El contenido que se busca no debe estar a más de tres clic de la página principal y la estructura de información no debe ramificarse mucho.

Las páginas son de tres tipos:

1. Noticias: Exigen actualizaciones en el contenido o información.
2. Filtros, índices o menús: Ofrecen información del contenido para decidir cuál explorar.
3. De destino o de contenido: Ofrecen los contenidos anunciados por las ligas. Estas, a su vez pueden ser:
  - ✓ Páginas principales: (dan información sobre el sitio, el autor, el contenido a través de índices o barras de navegación).
  - ✓ Páginas de contenido (ofrecen la información anunciada por los índices).
  - ✓ Páginas de links o enlaces (insertan enlaces activos a otras páginas o sitios de Internet).
  - ✓ Páginas de comunicación (ofrecen formatos para elaborar o responder preguntas).

### **Indicador de Utilidad de la información:**

Para que la página sea interesante, el usuario ha de encontrar utilidad en sus contenidos. Dichos contenidos han de servir para resolver determinadas necesidades de información de los usuarios. La información que coloque en su página debe responder a las necesidades del usuario.

## **2.2. Herramientas informáticas empleadas para desarrollar estudios webmétricos.**

Hoy en día, a través de Internet se realizan grandes estudios en los campos de la Cibermetría, siendo la Web, la principal herramienta informática para la obtención de la información y el conocimiento.

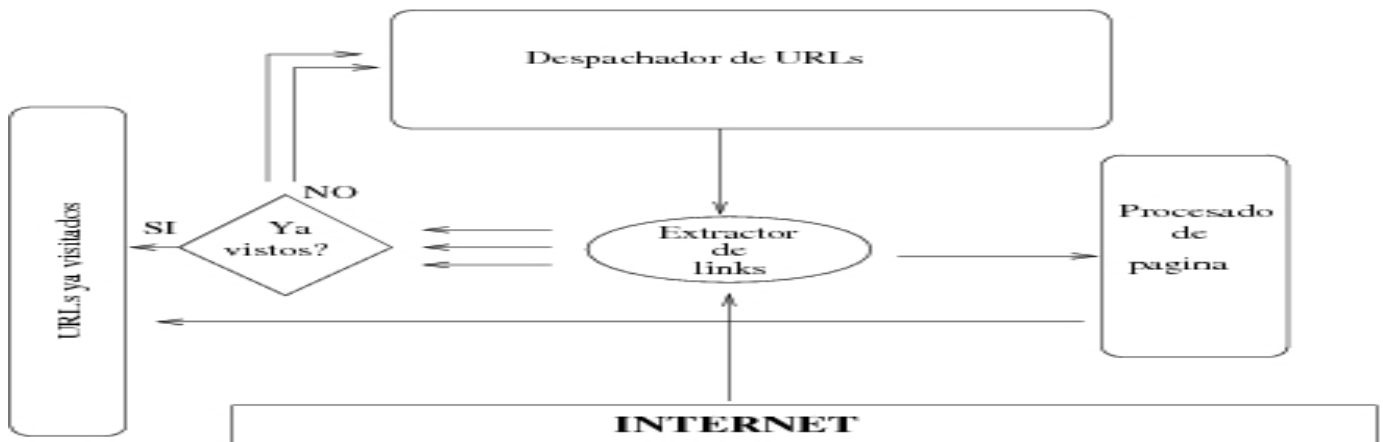
Las actividades que se realizan sobre la Web aumentan considerablemente cada día, como por ejemplo: la comunicación científica, la descripción y evaluación de los recursos web, el análisis cibernético, etcétera; siendo necesario cuantificar ciertos datos de interés para valorar la repercusión de estas actividades. Esto conlleva a utilizar nuevos grupos de herramientas webmétricas que permitan el tratamiento cuantitativo de los contenidos en la Web.

Los motores de búsquedas, también como herramientas sencillas y potentes para la búsqueda y extracción de datos cuantitativos de sedes web, proporcionan soluciones rápidas gracias a su spider (o Web Crawler) ya que estos permiten la indexación de los sitios web.

El spider es un rastreador web (también conocido como araña o robot web de Internet), el cual navega por la WWW de una manera metódica y automatizada para acceder a las páginas web en busca de contenido y enlaces. La visión del spider es únicamente de tipo texto, es decir, tanto las imágenes como contenidos flash o de tipo .OCX compilados y similares no son registrados ni escaneados por el buscador.

El procedimiento básico de un robot consiste en suministrar una URL<sup>3</sup> inicial o un conjunto de ellas, obtener la página web correspondiente y a continuación extraer todos los enlaces existentes en dicha página. Con los enlaces obtenidos es necesario realizar una serie de operaciones previas de normalización y luego comprobar las URL que se habían seguido previamente y, en caso de no haberlas recorrido, introducirlas en una cola de direcciones URL a seguir.

Como cada spider consta de una base de datos que sería un backup del código HTML de las páginas rastreadas, y estas se encargan de realizar sus propios esquemas de cada sitio web mediante índices, para aportar mayor rapidez al acceso de las páginas. La información se almacena en esta base de datos o bien en estructuras de ficheros con codificación ASCII; por último se obtiene la URL del siguiente enlace a seguir y comienza de nuevo el proceso. (Ver Figura 1)



**Figura 1: Esquema básico de un spider o robot.**

Un spider construye además, tablas y combinaciones de palabras cuya función es conseguir las mejores páginas y explorar los contenidos de todo el texto, de tal forma se evita tener que rastrear todas y cada

<sup>3</sup> Uniform Resource Locator. Indicador de la localización de un objeto en la Red Global Mundial.

una de las páginas del buscador. Otra función elemental de la que se encarga el spider es la de mostrar mediante un programa la visualización de todas las páginas indexadas, generando páginas en formato HTML con los resultados de la búsqueda que se haya realizado.

### 2.2.1. Herramientas Libres.

**WIRE Crawler.** (<http://www.cwr.cl/projects/WIRE/>): este proyecto chileno WIRE, es un esfuerzo iniciado por el Centro de Investigación Web (Center for Web Research) dirigido por el Dr. Ricardo Baeza-Yates, para crear una aplicación que permita la recuperación de información; diseñada para ser utilizada en la Web.

Actualmente el software WIRE incluye: un formato simple para almacenar una colección de documentos web, un rastreador web, herramientas para la extracción de las estadísticas de la colección y herramientas para la generación de informes acerca de la colección.

Las principales características del software WIRE son las siguientes:

- ✓ Escalabilidad: Diseñado para trabajar con grandes volúmenes de documentos, ha sido probado con varios millones de documentos.
- ✓ Prestaciones: Programado en C/C++ para un alto rendimiento.
- ✓ Configurable: Todos los parámetros para el rastreo y la indexación se pueden configurar a través de un archivo XML.
- ✓ Análisis: Incluye varias herramientas para analizar, extraer estadísticas y la generación de informes sobre sub-conjuntos de la Web, por ejemplo: la Web de un país o de una gran Intranet.
- ✓ De código abierto: El código está libremente disponible.
- ✓ Además el sistema está diseñado para centrarse en la evaluación de la calidad de la página, utilizando diferentes estrategias de rastreo y la generación de datos web para la caracterización de los estudios.

El robot WIRE se compone de diversos programas o módulos que lo ayudan con el funcionamiento, normalmente el mismo funciona de forma reiterativa, los programas son wire-bot-reset, wire-bot-seeder, wire-bot-manager, wire-bot-harvester, wire-bot-gatherer y wire-bot-run.

**CrawlTrack.** (<http://www.crawltrack.net/>): Desarrollado por Jean-Denis Brun y es una herramienta completa que ayuda a optimizar los sitios web, sirve para monitorear y rastrear la actividad de los motores de búsqueda en un sitio web. Es una gran pieza de software libre (licencia GNU GPL) y está disponible en una amplia gama de idiomas (existen por el momento en inglés, español, alemán, turco, neerlandés, ruso y francés).

CrawlTrack es un software de análisis web, que brinda información sobre el número y el origen de los visitantes, el porcentaje de abandonos, el número de errores 404<sup>4</sup>, el número de descargas de archivos, su uso de la carga del servidor y su indexación del sitio. Da a conocer las palabras claves utilizadas para encontrar el sitio, las páginas de entrada correspondiente, la tasa mundial de rebote, los navegadores utilizados por los visitantes del sitio, etcétera. Los datos se muestran en forma de tablas y de gráficos para una fácil visualización.

Muestra además, una vista de detalles de la actividad de orugas y los intentos de hacking, de esta forma ayuda a proteger los sitios web mediante el bloqueo de inyección de código o numerosas tentativas de inyección de SQL. Es muy utilizado para seguir a diario la posición que ocupa un sitio web en los motores de búsqueda y su rango para cada palabra clave que envían los visitantes.

Para su utilización sólo se necesita tener un host con PHP (versión > = 4.3.0 con GD2) y una base de datos MySQL. La instalación es completamente automática.

### 2.2.2. Herramientas Online.

**Emezeta Rank** (<http://rank.emezeta.com/>): Es un sistema online utilizado para estimar de manera rápida y sencilla la popularidad que tiene una página o sitio web en las diferentes redes de buscadores, y su éxito en referencias y enlaces. Brinda de cierta forma, las características que tiene una página web, sus puntos fuertes y débiles, dónde falla, y otros aspectos que posibilitan realizar una mejora sobre el sitio web.

Este servicio otorga además, una calificación simbólica a cada sitio web que se prueba, sumando ciertos valores como los enlaces recibidos, el PageRank o los resultados de la validación del código de la página. Brinda gran cantidad de información verdaderamente útil para el mantenimiento de cualquier sitio web,

---

<sup>4</sup> HTTP 404, error 404 o no encontrado es un código de estado HTTP, el cual indica que el navegador web ha sido capaz de comunicarse con el servidor, pero no existe el fichero que ha sido pedido.



entre las que se encuentran: las características del servidor y su instalación, valora los criterios de redirección WWW alternativos, la existencia de un sistema XML, comprueba la validación del fichero robots.txt, el código XHTML y el CSS.

**Bizinformation** (<http://bizinformation.org/cu/>): Es un servicio gratuito online de evaluación basado en la Web, que presta información sobre los sitios web, estadísticas de los visitantes y que fundamentalmente facilita una cotización estimada de cuánto cuesta un sitio web.

Para ello tiene en cuenta varios aspectos como: antigüedad del dominio, nombre y extensión, PageRank de Google, número de visitas y páginas vistas (no tiene información directa), principales fuentes de tráfico: buscadores, tráfico directo, referers (tampoco tiene información directa), ingresos por publicidad de Google AdSense (tampoco tiene información directa), suscriptores del blog (tienen información parcial), temática, autoridad y enlaces entrantes (backlinks), calidad del contenido y frecuencia de actualización, así como tipo de servidor y costes de mantenimiento estimados.

### **2.3. Selección de indicadores webmétricos y herramientas informáticas a utilizar para evaluar la calidad de los sitios web.**

Todo tipo de información es susceptible de ser evaluada, sobre todo si se requiere reunir una colección de utilidad para los usuarios. Puesto que la mayoría de la información es accesible a través de Internet mediante páginas web, al analizar la calidad de la información accesible en línea, también debe ser analizada la calidad de las páginas o sitios web que la soportan.

Es evidente que la evaluación de páginas o sitios web es necesaria por motivos cuantitativos y cualitativos, por el elevado número de páginas existentes y las base de datos de varios buscadores que ya han sobrepasado miles de millones de documentos publicados; siendo cada día más difícil lograr que páginas web se inserten en los buscadores y alcancen un adecuado posicionamiento en los mismos.

Por tanto, se requiere aplicar un proceso de evaluación a los sitios web, estableciendo criterios mediante el uso de indicadores webmétricos concretos y herramientas informáticas que sirvan de forma definitiva para el análisis de las páginas web.

### **2.3.1. Indicadores webmétricos seleccionados para el estudio.**

A continuación se da a conocer una serie de indicadores que pueden ser considerados como un índice de calidad de una página o sitio web.

#### **Indicador de factor de impacto web (WIF):**

Este indicador mide la relación existente entre el número de citas que recibe un sitio en forma de enlaces y el número total de páginas del mismo. Los dominios con mayor WIF son los más referenciados o citados, y por lo tanto, estos son los más conocidos.

La importancia de utilizar este indicador para que un sitio web sea conocido y recomendado mediante enlaces, viene dado por el incremento de su buena reputación a los ojos de los buscadores que la toman en consideración, para asignar sus recursos a la recolección e indexación de la información. Se establece también, con el propósito de percibir la visibilidad alcanzada por los sitios y la eventual relevancia del conocimiento que brindan a los cibernautas, dando así una idea del nivel de alcance de estos sitios.

#### **Indicador de visibilidad:**

Es un indicador que mide a grandes rasgos la presencia de un sitio web. Viene dado por el número de enlaces externos (de terceras sedes) recibidos por un dominio. Este atributo indica cuán "visible" o cuán bien posicionado se encuentra un sitio web, una vez mostradas las listas de salida de los motores de búsqueda cuando se consulta por temas relevantes a él.

Con este indicador se pretende resaltar la importancia con la que es percibido un sitio web por los usuarios, la ventaja que puede ofrecer frente a otros competidores su correcta aplicación, así como asegurar un caudal importante de nuevos visitantes de forma sostenida.

Resaltar que la visibilidad se relaciona también, con la calidad óptima de los enlaces externos que permiten el acceso a un sitio, así como el propio posicionamiento del sitio que contiene dichos enlaces, pues influye de manera significativa en el posicionamiento del sitio accedido mediante los links, haciéndolo de esta manera más visible.

#### **Indicador de popularidad:**

La popularidad de un sitio o página web está estrechamente vinculada al consumo de información del sitio. Esto puede medirse contando el número y describiendo las características de los visitantes y las

visitas que recibe el mismo. Por tanto, una página web es más importante a la hora de que los buscadores ordenan sus resultados, cuanto más enlaces en otras páginas indexadas le apunten. Es por esta razón que cuando se promociona un sitio, se procura que este sea conocido, es decir, se busca su popularidad a través de la indexación de varias páginas hacia este.

Hay que señalar sin embargo, que medir la popularidad de un sitio no es tan sencillo como se pretende ver en la explicación anteriormente dada. Ya que si el "índice de popularidad" fuera igual a la suma del número de enlaces que apuntan a un sitio, bastaría con crear una multitud de páginas que apunten a este sitio para mejorar su posicionamiento. Por tanto, existen algoritmos de popularidad que se defienden de esta posible trampa, de un modo sencillo pero efectivo.

El "índice de popularidad" es en realidad la suma del número de páginas que apuntan a un sitio ponderados por el índice de popularidad de cada una de ellas (medido en porcentaje sobre la máxima puntuación alcanzable). [20]

Este indicador se utiliza en el presente estudio, por la repercusión que puede aportar un sitio web popular cuando atrae hacia él la mayor audiencia de cibernautas posible. Siendo tanto así, que el simple hecho de que un sitio sea popular, ayuda también a estar mejor considerado entre la oferta de sitios de Internet, ya que los buscadores pueden admitir las páginas con mayor probabilidad y las pueden posicionar mejor entre los resultados. Este indicador además, sirve para premiar a las Webs de muchos contenidos, a través de los distintos enlaces que interconectan páginas cargadas de información valiosa para los usuarios.

### **Indicador de accesibilidad:**

El indicador de accesibilidad define la posibilidad de que un sitio web pueda ser accedido y usado por el mayor número posible de usuarios. Su concepción está enfocada al acceso pleno del sitio, independientemente de las limitaciones propias de cada individuo, de las características de su equipo de navegación o del entorno ambiental desde donde se accede a la Web y de las derivadas del contexto de uso; posibilitando así, la realización de las tareas por parte de los usuarios, sin limitación alguna.

Se establece este indicador fundamentalmente, con el objetivo de aumentar el número de visitantes al sitio, mejorar la indexación de las páginas web en los buscadores y lograr además la fácil navegación del mismo. Para este último, se tienen en cuenta entre otros aspectos, la calidad de los enlaces tanto

externos como internos, los cuales deben permanecer constantemente actualizados para permitir de forma fácil y rápida el acceso a la información deseada.

Cuando se desea que la información sea accedida por todos, siempre se debe pensar en que la audiencia de la Web es realmente global, y no local. Por tanto, este indicador es muy importante para aquellas personas de edad avanzada con dificultades producida por el envejecimiento, para los usuarios con falta de recursos para acceder a los servicios de Internet, para los que no dominan el idioma, como es el caso de los de habla extranjera o los de menor nivel cultural, y también para los usuarios inexpertos, entre muchos más.

### **Indicador de contenido:**

Posiblemente es el indicador más importante a tener en cuenta para que un sitio logre una gran presencia en la Web. De la información suficiente, valiosa y actualizada, así como del lenguaje mayoritario de dicho contenido, dependen en gran medida que un sitio logre ser interconectado a través de enlaces por muchas otras páginas web. De esta forma el contenido de un sitio será más conocido, alcanzando así mayor visibilidad.

Este indicador se utiliza en el estudio, por el inmenso aporte que brinda a los buscadores, pues los favorece de un mayor volumen de información y variedad de contenidos que se pondrán a disposición de los usuarios. Además, conocer que estos motores de búsqueda indexan las páginas web basándose en el contenido de estas a partir de las palabras claves, permitiendo determinar la relevancia del contenido si dicha página web ha alcanzado una posición notable en los resultados de la búsqueda.

Permite medir el número y el tamaño de los objetos informáticos encontrados en cada uno de los sitios web, tal como los llamados ficheros ricos (ppt, pdf, doc, pps), ficheros de texto y las páginas dinámicas.

### **Indicador de posicionamiento:**

Este indicador tiene que ver con un grupo de técnicas y procedimientos que posibilitan colocar un sitio web en el mejor lugar, dentro de los resultados proporcionados por un motor de búsqueda o buscador. Por esta razón se debe tener en cuenta que, una página web bien posicionada se encuentra antes, se visita más, será conocida por los usuarios y si su contenido les resulta interesante entonces establecerán enlaces a dicha página contribuyendo a elevar su posición.

Por esta causa, se establece el posicionamiento de páginas o sitios web en los motores de búsqueda, como un indicador principal en la recuperación o publicación de la información en Internet, y por ende, para conseguir una adecuada presencia en la Web, ya que si estas páginas web no presentan una buena posición, entonces su información no será tan visible como para ser accedida por parte de los usuarios. Además, la selección de este indicador se ha establecido principalmente partiendo de que, una posición relativa puede ser un buen punto de análisis para medir la visibilidad y el impacto de las sedes o sitios web, y de esta forma valorar su aportación e importancia respectivamente.

### **2.3.2. Herramientas informáticas seleccionadas para el estudio.**

Por adaptarse a las necesidades del estudio de la valoración de la calidad de los sitios web, una de las herramientas estadísticas seleccionada es el **WIRE Crawler**. Además, el software presenta ciertos requisitos fundamentales como la velocidad, la eficiencia y la robustez, necesarios para el estudio que se quiere llevar a cabo. Y los programas para la recuperación de información de la Web que lo componen. Está disponible bajo licencia GPL.

La herramienta genera estadísticas importantes y varios reportes imprescindibles para confeccionar un estudio de carácter cualitativo sobre una Web analizada. Está diseñado para trabajar con grandes volúmenes de información; está programado en C++ para un alto rendimiento y es altamente configurable, pues todos los parámetros para el rastreo, la indexación, el análisis de los datos y la creación de los reportes estadísticos, se pueden configurar a través de un archivo XML.

Por otro lado, incluye varias herramientas para analizar, extraer estadísticas y generar una serie de informes sobre subconjuntos de la Web.

La aplicación genera 6 informes sobre la colección:

- ✓ Informe general sobre la colección descargada.
- ✓ Informe sobre las extensiones encontradas en el recorrido.
- ✓ Informe sobre los sitios web analizados en el recorrido.
- ✓ Informe sobre los enlaces de los sitios web analizados en el recorrido.
- ✓ Informe sobre los idiomas o lenguas encontrados en el recorrido.
- ✓ Informe sobre los distintos ciclos realizados por el Módulo de Recolección de Datos.

Se utiliza el **Web Link Validator** para la verificación de enlaces en sitios web, por los beneficios que su uso proporciona, ya que no solo se limita al control de los vínculos, sino también al problema de los archivos huérfanos y los errores en el código HTML. A través de sus filtros se descubren nuevos problemas como la lentitud de la carga, la existencia de páginas pequeñas, nuevas y viejas. También monitorea los cambios de las páginas web y archivos especificados. Es altamente personalizable y seguro. Protege además contra las sustituciones indebidas mediante la detección de vínculos redirigidos.

Se utiliza también el servicio online **Bizinformation** como herramienta para determinar, en este caso, el total de páginas que posee un determinado sitio web, así como un aproximado del total de visitas que reciben estos sitios web diariamente. Aunque su objetivo principal es calcular cuánto vale el sitio web como ya se ha planteado en epígrafes anteriores, su función en la investigación no será esta; sino la de aportar solamente, los parámetros a tener en cuenta para el cálculo del costo. De esta manera se puede comprobar el consumo de información que de estos sitios web se tiene, y por lo tanto contribuir a medir posteriormente el indicador de contenido.

Se establece el sistema online **Emezeta Rank** como una herramienta estadística cuyo principal objetivo es brindarle a los diseñadores de sitios web, así como a los webmáster, un índice de popularidad válido que permita determinar cuán popular resulta el sitio. Posibilita además medir el comportamiento de los indicadores de visibilidad y posicionamiento web a través de varios puntos de análisis. Esta herramienta, al proporcionar los índices numéricos relacionados a la cantidad de enlaces externos provenientes de Google, Yahoo y Live unido a las referencias en Google Blog Search, permite analizar el indicador de visibilidad estableciendo un promedio entre estos índices.

Emezeta Rank también proporciona el **PageRank de Google**. Este no es más que un número del cero al diez (máximo PageRank posible) y que Google utiliza para determinar la importancia de una página web. Este es uno de los muchos factores utilizados para determinar las páginas que aparecen en los resultados de búsqueda. Son muy pocos los sitios que gozan de una máxima calificación; 1 es la menor calificación que puede recibir un sitio normal, y cero significa que el sitio ha sido penalizado o aún no ha recibido una calificación de PageRank.

La teoría del PageRank se basa en que los sitios web más importantes son aquellos que tienen la mayor cantidad de enlaces entrantes, pues en Internet se consideran a estos enlaces como un voto a favor del

sitio hacia el cual apuntan. Dicho "voto" tiene un peso dependiendo de la importancia del sitio en que se encuentra el enlace; es decir, los enlaces que salen de los sitios más importantes tienen mayor peso y viceversa. Pero resulta que, el peso de votación de una página web se divide equitativamente entre todos los enlaces de esta, lo cual trae consigo que mientras más enlaces hay en dicha página menos peso tiene el voto de cada enlace. Razón por la cual se toma este PageRank, como dato relevante para medir el factor de impacto web, necesario para los fines del estudio.

Considerando que el PageRank podría fallar si fuese el único factor involucrado en posicionar los sitios web, ya que puede ser manipulado mediante campañas de mercadeo electrónico honestas o fraudulentas, **Google** toma en cuenta para el posicionamiento otros factores de gran importancia. Utiliza el PageRank como un indicador del alcance de un sitio, pero además combina esta información con otra serie de análisis para determinar la relevancia de los resultados.

Google analiza también el contenido del sitio diseñado, así como el contenido, importancia y reputación de los sitios que apuntan hacia este. Además, toma en consideración las palabras claves que se encuentran en los sitios que apuntan al sitio web analizado para determinar si los enlaces que se reciben tienen valor. Todo lo planteado da la medida de que para el estudio en cuestión, el PageRank de Google se establece como apoyo para el análisis del indicador de posicionamiento.

En el estudio se hace uso del buscador **Alexa** para analizar además al indicador de popularidad, ya que este buscador intercepta visitas en todo el mundo y establece a partir de estas un Ranking de popularidad. Se utiliza por la facilidad de lograr la medición de este indicador a partir de los valores relativos proporcionados por dicho buscador.

Alexa ofrece estadísticas amplias sobre cualquier página web y ciertas variables a considerar para el posicionamiento web. Dentro de estas estadísticas en "Search Analytics", se visualiza una tabla con las búsquedas realizadas a una determinada web, su factor de impacto, popularidad y las palabras claves.

Otro resultado que brinda Alexa es un informe sobre los sitios web que reciben más tráfico en una búsqueda realizada y el porcentaje de importancia de cada sitio. Los archivos de Alexa contienen la mayor colección de datos de la Red de Internet en todo el mundo, más que Google y más que cuatro veces AltaVista.

Vale destacar que, la posición que ocupe un sitio web en el Ranking de este buscador depende del tráfico que cada uno de estos sitio web reciba de personas cuyo navegador (solo iExplorer o Firefox) tenga instalada la barra de Alexa.

Lograr un entendimiento de cómo se manifiestan los indicadores webmétricos en los distintos sitios web, ha sido la meta de este capítulo, en el cual, se definen los diferentes indicadores que repercuten en la calidad de un sitio web así como las herramientas webmétricas e informáticas que visualizan la presencia de los mismos en una aplicación web.

El estudio se centra a continuación en procesar todos los datos recolectados y documentar cuantitativamente como influyen dichos indicadores en la calidad de un sitio web; de manera que ayude a mejorar específicamente la calidad de los sitios web de la UCI. Se establece entonces, una estrategia metodológica que guie la usabilidad de estos indicadores.



## Capítulo 3: Estrategia metodológica para mejorar la calidad de los sitios web en la UCI.

En el presente capítulo se brindan los resultados alcanzados del estudio webmétrico realizado sobre la calidad web de los sitios web, a través del uso de indicadores webmétricos que influyen en esta. Se elabora además una estrategia metodológica a seguir para mejorar paulatinamente la calidad de los sitios web de la UCI a través del uso de indicadores webmétricos.

### 3.1. Fundamentación de la estrategia.

La estrategia metodológica que se propone, sobre la base de los indicadores webmétricos, tiene como **objetivo**: mejorar la calidad de los sitios web en la UCI y está orientada hacia el reforzamiento de estos.

De modo que el diseño de la estrategia se sustenta en concepciones de investigadores cubanos y extranjeros, las cuales se complementan y, resultan ser lo suficientemente generales y flexibles para satisfacer los requerimientos de la misma. Dicha concepción se refiere a la transformación del estado real al estado deseado del objeto en transformación, lo que se ha de expresar en un sistema de pasos o acciones para lograr el fin esperado.

Tales acciones y pasos están determinados por los indicadores que permiten medir la calidad de los sitios web, requiriendo una gran implicación y compromiso por parte de los actores involucrados.

Es preciso declarar que para la estrategia se distinguen dos componentes principales: metodológicos y prácticos, consta de dos fases: el diagnóstico y la intervención. En su estructura tanto los elementos metodológicos como los prácticos están interrelacionados de forma sistémica. Los prácticos persiguen un objetivo específico y están constituidos por un sistema de acciones que, para su ejecución se basan en las orientaciones de los componentes metodológicos. Todos, de conjunto, tributan a los objetivos generales de la estrategia.

De modo que por su carácter general y flexible, ambos tipos de componentes están concebidos no de una manera única, para ser aplicados de una vez y por todas, sino como un proceso que comienza por la adecuación y contextualización a las condiciones concretas de la institución en estudio, con la participación de los involucrados. Por ello, requiere imprescindiblemente que se complementen con la labor especializada de cada miembro en la elaboración de las páginas web.

### 3.1.1. Componentes metodológicos de la estrategia.

Se elaboraron a partir de los estudios webmétricos realizados sobre la calidad de los sitios web, a través del uso de indicadores webmétricos que influyen en dicha calidad. Estos estudios se efectúan primeramente, sobre una muestra significativa de los sitios web representativos de las Universidades Latinoamericanas según el Ranking Web de Universidades del Mundo de enero de 2010, y luego sobre una segunda muestra correspondiente a los sitios de la UCI.

De igual manera, se establecen los resultados obtenidos de los sitios representativos de las Universidades Latinoamericanas, que se encuentran situados en una Web visible<sup>5</sup> y los sitios de la UCI que están presente en una Web invisible<sup>6</sup>. Esto se efectúa con el objetivo de observar el comportamiento de los indicadores webmétricos en sitios web presentes en Webs con diferentes características (visible e invisible), de forma tal que se obtenga cómo estos indicadores contribuyen a la calidad de los sitios web y cómo influye el estar posicionado en dicha calidad.

A partir de este análisis, se valoran aquellos indicadores ausentes en los sitios web de la UCI necesarios para adquirir una buena calidad del sitio, y se propone una metodología a seguir para mejorar paulatinamente la calidad de los sitios web de la UCI a través del uso de indicadores webmétricos.

Los componentes metodológicos no están concebidos para incorporarlos total o parcialmente de manera necesaria a los sitios web, o de la gestión institucional, sino para que se les considere en la elaboración y el perfeccionamiento de estos sitios. Son una contribución para el perfeccionamiento, pero requiere de la intervención de cada componente del equipo de trabajo y lejos de intentar ser una indicación, constituye una guía para la acción.

Los componentes metodológicos se conformaron a partir de un **problema científico**, en este caso: Cómo mejorar la calidad de los sitios web en la UCI? Es por ello que la solución del problema precisa de la implementación de la estrategia, y por lo tanto, alcance un índice de calidad, desarrollo y operatividad del

---

<sup>5</sup> Comprende sitios cuya información puede ser indexada por los robots de los buscadores convencionales y recuperada casi en su totalidad mediante una consulta a sus formularios de búsqueda.

<sup>6</sup> Comprende sitios cuya información, generalmente es almacenada y accesible mediante bases de datos. Son páginas que no tienen una URL fija y que se construyen en el mismo instante (temporales), desapareciendo una vez cerrada la consulta.

recurso tecnológico e impacto notable en el campo de estudio al que pertenece, capaz de transformar su realidad desde cualquier ámbito.

A los efectos de esta estrategia, se considera como un concepto operacional el término: “indicadores webmétricos”, un tipo de generalización y conocimientos informáticos que orientan y sirven como tabla para medir la calidad de los sitios web.

Al responder a condiciones socioeconómicas, dentro de su estabilidad y carácter permanente, son flexibles, cambiantes y no acabados; se enriquecen, perfilan y modifican a través de la dialéctica de las ciencias informáticas, pero sí son condición y requerimiento para perfeccionar los sitios web en la UCI.

Así pues, los indicadores webmétricos constituyen una importante herramienta metodológica para la aplicación de la estrategia, en tanto orientan el enfoque del contenido, métodos y las acciones de transformación. Estos indicadores son: visibilidad, popularidad, factor de impacto web, accesibilidad, contenido y posicionamiento web.

### **3.1.2. Métodos y procedimientos utilizados para la elaboración de la estrategia:**

#### **Observación científica:** *(Ver Anexo 1)*

Conocido el problema, mediante el cual se conformaron los componentes metodológicos de la estrategia, y el objeto de investigación, los sitios web, se estudió el curso natural del comportamiento de los indicadores webmétricos en los sitios web correspondientes a las muestras de estudio. Esto fue realizado sin alterar las condiciones naturales, es decir, simplemente se contempló la presencia de indicadores para configurar una base de conocimiento fiable, sobre la influencia de estos en la calidad de los sitios web.

Durante la observación se utilizaron varias tecnologías y herramientas informáticas que posibilitaron la realización de este proceso, entre las que se encuentran:

- ✓ Sistema Operativo: Ubuntu GNU/Linux.
- ✓ Sistema Operativo: Windows XP, versión 2002.
- ✓ Spider para la Recolección de Datos: WIRE Crawler Software V.0.21.
- ✓ Buscador Alexa.
- ✓ Web Link Validator. V 5.1.
- ✓ Sistema online Emezeta Rank.

- ✓ Sistema online BizInformation.

**Selección de la muestra:**

Para la selección de las muestras de estudio se utilizó un tipo de muestreo no probabilístico, específicamente el muestreo por cuotas. En este tipo de muestreo se fijan unas "cuotas" que consisten, en este caso, en un número de sitios web que reúnen determinadas condiciones desde el punto de vista webmétrico. Dentro de estas se destacan la presencia de los diferentes indicadores ya mencionados: visibilidad, accesibilidad, popularidad, contenido, posicionamiento y factor de impacto web. Luego una vez determinadas las cuotas se eligen los sitios web que se encuentren en las poblaciones y que cumplan estas condiciones, conformando de esta manera las muestras del estudio, que no son más que, un subconjunto representativo de dichas poblaciones en las que la presencia de los diferentes indicadores webmétricos influye sobre la calidad de estos sitios web.

Para realizar el cálculo de la muestra representativa de las poblaciones del estudio, se toman en cuenta las siguientes fórmulas: [21]

$$n = \frac{n'}{1 + n' / N}$$

$$n' = \frac{Z^2 (p) (1-p)}{e^2}$$

Donde: **n'** representa la cantidad teórica de elementos de la muestra; **n** denota la cantidad real de elementos de la muestra a partir de la población asumida; **Z** se corresponde con el nivel de confianza; **N** representa el número total de sitios web que conforman la población; **p** denota la proporción esperada o estimada de la muestra y **e** no es más que el porcentaje de error.

Se obtienen de esta manera **dos muestras: n1**, conformada por 63 sitios web procedentes de la primera población con un nivel de confianza (Z) de 2.58 puesto que se considera realizar el estudio con un 99% de confianza, una proporción esperada (p) de 0.94 y un porcentaje de error (e) de 5% para una población de 100 sitios web (N); y **n2**, conformada por 113 sitios web procedentes de la segunda población con un nivel de confianza (Z) de 1.96 puesto se considera realizar el estudio con un 95% de confianza, una proporción esperada (p) de 0.89 y un porcentaje de error (e) de 4% para una población de 219 sitios web (N).

### **3.1.3. Establecimiento de principios o criterios de evaluación.**

Para medir la calidad de un determinado sitio web, se conforman, en este caso, criterios de evaluación que incluyen un conjunto de aspectos a observar, de manera que se determine su grado de cumplimiento.

Los criterios de evaluación, que coinciden con los indicadores webmétricos definidos son los siguientes:

- ✓ Accesibilidad: entendida como la capacidad del sitio web para estar al alcance de todos los usuarios.
- ✓ Contenido: entendido como la disposición del sitio web para proveer información relevante.
- ✓ Popularidad: entendida como la cualidad del sitio web de lograr una proyección en su imagen de manera correcta y adecuada.
- ✓ Impacto: entendido como la amplitud con la que se ha dado a conocer el sitio web en forma de enlaces.
- ✓ Visibilidad: entendida como la utilidad del sitio web de ser referenciado por muchas otras páginas web.
- ✓ Posicionamiento: entendido como la habilidad del sitio web de estar entre los primeros lugares de los resultados mostrados por los buscadores.

### **3.2. Análisis de los sitios web representativos de: las Universidades Latinoamericanas según el Ranking Web de Universidades del Mundo y los sitios web de la UCI.**

Como primera fase de la estrategia es preciso realizar un diagnóstico para identificar la situación actual de las páginas web de las Universidades Latinoamericanas y de la UCI. De esta forma se extraen los datos que permiten cuantificar la presencia de los indicadores webmétricos en dichos sitios web, con el fin de valorar la calidad de estos.

#### **3.2.1. Evaluación de los sitios web correspondientes a las Universidades latinoamericanas.**

El Ranking Web de Universidades del Mundo de enero de 2010, está basado en indicadores web y este refleja mejor la actividad global de las instituciones a diferencia de otros Rankings. El mismo a su vez, constituye un indicador de impacto, rendimiento y prestigio para las instituciones educativas en todo el mundo.

En el estudio de esta primera muestra se considera el análisis cuantitativo que el mismo Ranking ofrece, además de un grupo de indicadores webmétricos como ya se ha mencionado, para evaluar la calidad y la persistencia de estos sitios en la Web.

Para analizar cada uno de los sitios web de la primera muestra de estudio, atendiendo a la presencia de los indicadores webmétricos definidos con anterioridad, se establece la siguiente interpretación general a la puntuación que de los indicadores se conceda más adelante, una vez hayan sido analizados por las herramientas informáticas determinadas anteriormente.

Puntuación.	Interpretación a la presencia de los indicadores.
10	Excelente
6-9	Buena, pero mejorable
4-5	Regular
1-3	Poca
0	Ninguna

**Tabla 2: Interpretación general a la puntuación de los indicadores.**

La puntuación total para un sitio web se obtiene entonces, dividiendo el total de puntos alcanzados por cada sitio web entre el número de indicadores que se estudian. Por tanto: Si el sitio obtiene una puntuación de 10 puntos, entonces se valora de excelente su calidad en cuanto a la aplicación de estos indicadores webmétricos. Por lo que se deduce que entre 7 y 9 puntos la calidad del sitio web es buena, entre 5 y 6 suficiente. Por último, menor de 5 insuficiente.

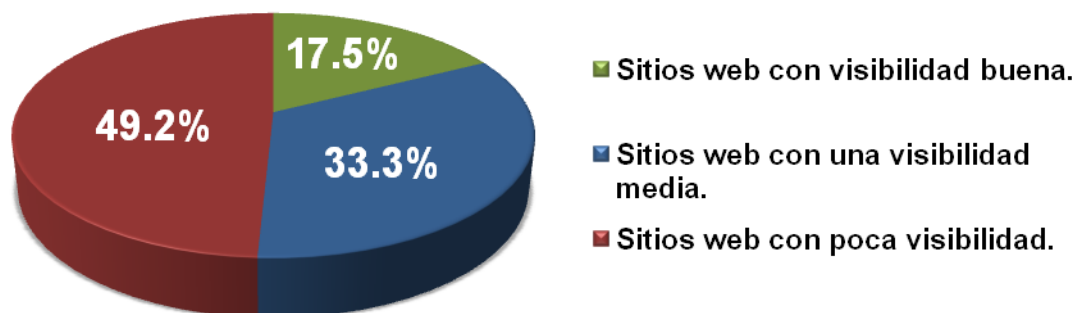
El rango de puntuación fijado de 0 a 10, no se obtiene de la misma manera para cada uno de los indicadores a identificar en estos sitios web. Por tanto, a continuación se muestra cómo se puntea cada indicador y además se presentan los resultados alcanzados en cada análisis:

### **Indicador de visibilidad:**

Como se viene explicando, la visibilidad de un sitio web se encuentra relacionada con la cantidad de enlaces que recibe este, provenientes desde otros sitios o páginas web, díganse los enlaces externos. Para determinar la presencia de este indicador se establece la puntuación del promedio calculado, referente al tema en cuestión, a partir de los valores ofrecidos por los buscadores Google, Yahoo, Live y

las referencias en Google Blog Search, todos proporcionadas en base a 10, a partir del sistema online Emezeta Rank.

El siguiente gráfico muestra cuán percibido o visible son los sitios analizados en la Web, por los usuarios de Internet.



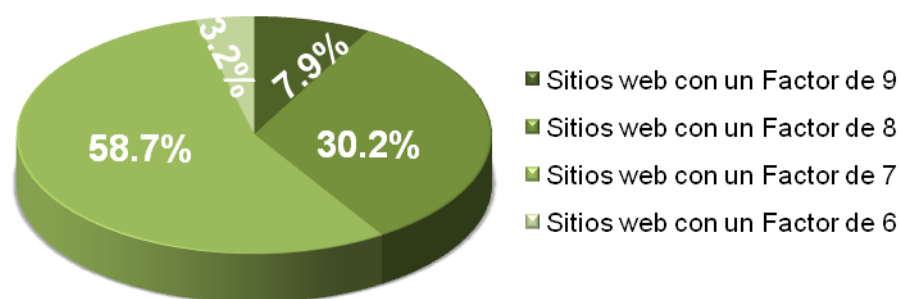
**Gráfica 1: Presencia del indicador de visibilidad en los sitios web latinoamericanos.**

Los resultados obtenidos evidencian que predominan los sitios web con mayor visibilidad hipertextual de sus páginas en los distintos buscadores, específicamente un 50.8%, lo cual implica que la calidad de los enlaces externos mediante los cuales se pueden acceder a los sitios web representativos de estas Universidades es notable, y esto influye a su vez de manera significativa en el propio posicionamiento del sitio por ser más visible.

**Indicador de factor de impacto web:**

En el pasado se utilizaba el llamado WIF, que se obtiene como cociente del número de enlaces entre el número de páginas de una sede o de un dominio. Por distintas razones, este índice da lugar a numerosos artefactos matemáticos, por lo que ha dejado de utilizarse. Una alternativa, bastante difícil de calcular, es el índice que se construye de acuerdo al peso relativo de las sedes de origen de los enlaces, más conocido como el PageRank de Google, pues su teoría se basa en que los sitios más importantes son aquellos que tienen la mayor cantidad de enlaces entrantes, enlaces considerados como un voto a favor del sitio hacia el cual apuntan. Esto ratifica que los dominios con mayor WIF son los más referenciados o citados, y por lo tanto, los más conocidos.

Para el estudio en cuestión, este indicador se obtiene mediante el sistema online Emezeta Rank, (PageRank proporcionado en un rango de números enteros entre 0 y 10). La siguiente gráfica muestra cómo se evidencia este indicador:



**Gráfica 2: Presencia del indicador de factor impacto web en los sitios web latinoamericanos.**

En término general se puede decir que, todos los sitios web de esta muestra de estudio según dicho PageRank, presentan un WIF por encima de los 5 puntos, lo cual indica que son bien conocidos por la mayoría de los internautas. Esto viene dado por la gran cantidad de sitios web que los referencian, facilitando que proporcionen descripciones exactas de la estructura y actividades de las universidades que representan. De esta manera les permite atraer nuevos estudiantes y académicos de todo el mundo.

#### **Indicador de popularidad:**

El presente indicador, que permite determinar cuán conocido o visible se ha convertido un sitio web en la Red de Internet, se cuantifica a través de la posición de popularidad que ocupa cualquier sitio web en el Ranking de popularidad que determina el buscador Alexa. Para ello, según lo publicado en el artículo: *“Sobre la pertinencia de Alexa, sus números y algunas ideas para mejorarlos en su web”*, [22] se asumen los mismos cuatro niveles significativos que este plantea para analizar el valor de la popularidad asignado a las posiciones del Ranking de Alexa y de esta manera comprender cuán precisos son sus datos, es decir, cuán importante es la diferencia entre la posición de un sitio web contra otro.

El primer nivel proporcionado por este artículo se refiere a las posiciones entre 1 y 20.000, quizás 50.000. Luego están los sitios con las posiciones entre 50.000 y 100.000, cuyas estadísticas son, aparentemente, muy aceptables. Se estima además que un tercer nivel de precisión está en las posiciones entre 100 mil y, quizás, 300 mil o más, específicamente 500.000. Destacar para este nivel que por encima de 100.000, ya



Alexa no ofrece muchos detalles, que sí son indicados por debajo de ese número, de todos modos, el cuarto nivel de precisión se coloca, por comparaciones entre sitios web, en las posiciones comprendidas desde 501.000 hasta un millón.

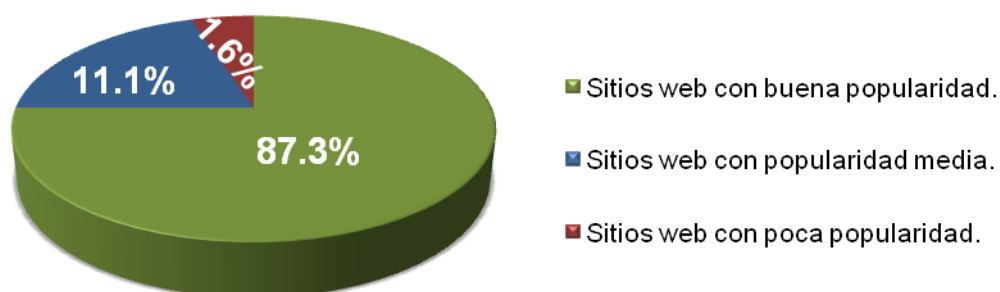
En la siguiente tabla se recogen estos niveles significativos y son desglosados para hacerlos coincidir apropiadamente a la interpretación de la *Tabla 2*, relacionada a la presencia del indicador en cuestión en un rango de valores entre 1 y 10.

Posición	< 20000	20001-40000	40001-60000	60001-80000	80001-100000
Puntos	10	9	8	7	6

100001-200000	200001-300000	300001-400000	400001-500000	> 500000
5	4	3	2	1

**Tabla 3: Puntuación asociada a las posiciones adquiridas a través del buscador Alexa.**

Una vez que se ajustan estos valores a los sitios web analizados, se obtiene que un 87.3% de estos sitios presentan una buena popularidad en la Web. Esto permite asentar que dichos sitios son bien considerados entre la demanda de sitios de Internet, así como premian a esta Web de variados contenidos de información. La gráfica mostrada a continuación permite comprender la presencia de este indicador en la muestra de estudio:



**Gráfica 1: Presencia del indicador de popularidad en los sitios web latinoamericanos.**

### Indicador de accesibilidad:

Como se viene analizando, la calidad de los enlaces de un sitio web posibilita en gran medida que dicho sitio sea de fácil indexación para los buscadores y por lo tanto más accesible por todos. Para determinar la

presencia de este indicador, se obtiene, a partir de la herramienta Web Link Validator, un porcentaje de calidad de enlaces buenos así como rotos. Dichos porcentajes se proporcionan de 0 a 100%, y se les ha hecho corresponder en rangos, un equivalente del 1 al 10 para de esta manera, vincularlo a la interpretación de la *Tabla 2*.

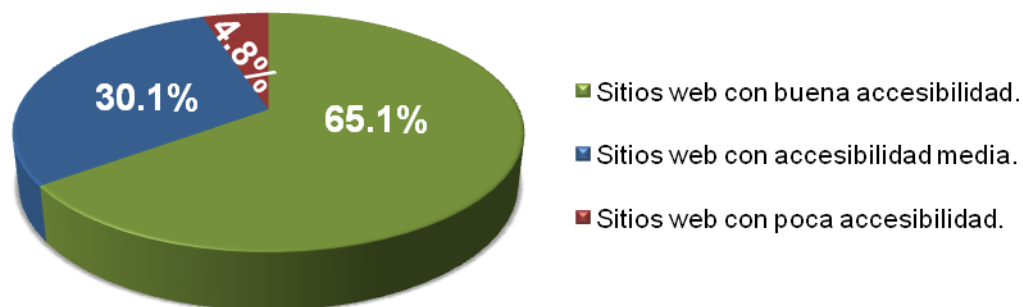
Calidad Enlaces(%)	1-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80	81-90	91-100
Puntos	1	2	3	4	5	6	7	8	9	10

**Tabla 4: Puntuación asociada al porcentaje de calidad de los enlaces.**

A esta valoración se le incorpora también otro criterio, y este se relaciona con el lenguaje mayoritario por sitio web, o sea, la diversidad de idiomas contribuye a que un mayor número de usuarios pueda acceder a la información que brinda el sitio. De ser brindada la misma en 5 idiomas o más, se le asigna un valor de 10, la presencia de 4 idiomas implica un valor de 8 y solo 3, 2 ó 1 idioma(s), los valores de 6, 4, y 2 respectivamente.

Llegado hasta aquí se estima entonces, un valor promedio de accesibilidad entre 1 y 10, que contribuya a la interpretación descrita en la *Tabla 2*.

La siguiente gráfica muestra como se comporta este indicador en los distintos sitios web que se analizan a partir de los aspectos mencionados.



**Gráfica 4: Presencia del indicador de accesibilidad en los sitios web latinoamericanos.**

Una vez que se analiza la gráfica permite comprender que una buena calidad de enlaces, tanto internos como externos, así como lograr una traducción de los principales servicios que brinda el sitio web incluidos documentos científicos o de otra índole, contribuye a aumentar el número de visitantes al sitio,

mejorar la indexación de las páginas web en los distintos buscadores, además de lograr una fácil navegación dentro del mismo.

### **Indicador de contenido:**

El contenido de los sitios web, no es más que la información que se expone en la Red. Por ello es de vital importancia que sea relevante y accesible a todos, para de esta manera conseguir Rankings altos en los buscadores sin caer en la penalización por contenido duplicado. Por tal razón, este indicador se va a analizar atendiendo entonces, al cálculo del promedio entre el número o tamaño de la información que aportan a la Web estas Universidades, así como al consumo de información que en estas se presenta y que se basa en conceptos no tan habituales como uso de información o comportamiento de la información.

El primer aspecto que se plantea para determinar la presencia del indicador en cuestión, está relacionado con el número de documentos, y este a su vez, es medido a través del total de ficheros ricos que contiene un dominio web. El segundo está vinculado al tráfico de información que de los sitios se obtenga, puesto que a mayor tráfico mayor consumo.

Por esta razón para determinar cuánto tráfico de información poseen los sitios web que se analizan, en el presente estudio se considera lo planteado en el artículo: *"Garantice un aumento del tráfico en su sitio web"*, [23] en el cual se establece que un tráfico elevado depende, entre otros parámetros, del número total de páginas que tiene un sitio y la cantidad de visitas que este recibe.

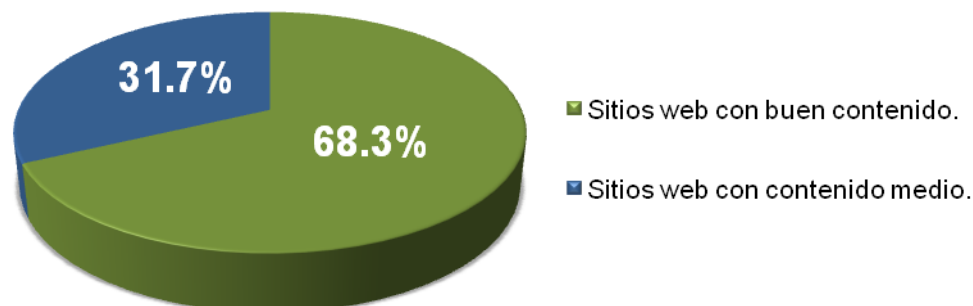
Además se plantea: *si se tiene un sitio web compuesto por 10 páginas solamente y cada página recoge una media de 7 visitantes por día, esto significa 70 visitas al día aproximadamente, lo que no descarta la posibilidad de que hayan menos, (10 páginas x 7 visitantes por página). Si ahora se contaran con 100 páginas similares a las 10 con las que se contaba antes, se tendría aproximadamente  $7 \times 100 = 700$  visitas al día en vez de 70. Lo que ratifica también que mientras se posean más páginas, se obtiene más tráfico.* [23]

El proceso de cuantificación para el caso del número o tamaño de la información que aportan a la Web estos sitios a través de los ficheros ricos, se establece a partir de la recuperación directa a través del

motor de búsqueda Google según la sintaxis **site:** *url del sitio* **filetype:** (*doc, pdf, ppt*), que brinda la cantidad aproximada de cada uno de estos tipos de ficheros en el sitio.

Para determinar los valores del otro parámetro, o sea, el tráfico de información del sitio, se hace uso el sistema online Bizinformation, ya que el mismo ofrece los aspectos que se tienen en cuenta para analizar dicho tráfico, según lo que se viene planteado acerca de él.

Como la obtención de los datos para estos dos parámetros no se encuentra sobre una escala del 0 al 10, se hizo necesario establecer una analogía al valor correspondiente de cada aspecto con esta escala. Para ello se calculó su logaritmo (log), tal y como se describe en el artículo: “*Calidad de la información web en la banca electrónica*”, [24] puesto que esta transformación logarítmica permite hacer comparaciones a una escala más pequeña. Por ejemplo:  $\log_{10}(100)=2$  y  $\log_{10}(100000)=5$ . Es más fácil observar la diferencia entre 5 y 2 que entre 100000 y 100. Una vez determinada esta normalización logarítmica, se divide la suma obtenida entre 2, y se obtiene así, una valoración entre 1 y 10 para este indicador, mostrándose a continuación su comportamiento:



**Gráfica 5: Presencia del indicador de contenido en los sitios web latinoamericanos.**

Los resultados evidencian que estas Universidades ofrecen suficientes contenidos actualizados, dado por el elevado consumo de información que de estos sitios se adquiere.

**Indicador de posicionamiento:**

El presente indicador, como se viene analizando permite, entre otros criterios, lograr que un sitio web alcance una adecuada presencia en la Web. En este estudio, se tiene en cuenta para analizar dicho indicador, la propia posición que ofrece el Ranking Web de Universidades del Mundo para los sitios web

representativos de las Universidades de Latinoamérica, puesto que el mismo está basado en un conjunto de indicadores webmétricos, necesarios para el propio posicionamiento de estos sitios como se ha planteado en epígrafes anteriores, además de valorar los criterios relacionados con el PageRank que Google ofrece. Llegado hasta aquí se establece entonces, un promedio entre estos dos parámetros.

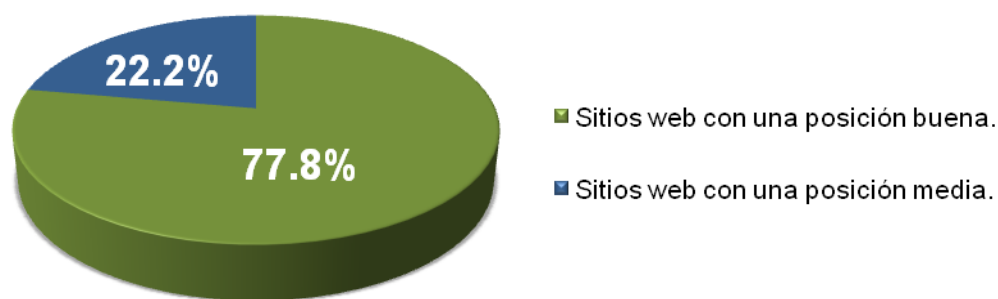
Debido a que Google brinda los valores de este PageRank a través de una escala del 0 al 10, para los fines de la investigación se hace necesario establecer una puntuación que esté acorde a la posición relativa adquirida por estos sitios web según dicho Ranking Latinoamericano (Top de 100 Universidades) y que haga más fácil la determinación de la presencia de este indicador a partir del promedio ya planteado.

La siguiente tabla, muestra la asignación de los valores del 1 al 10 para las posiciones alcanzadas por los sitios web de las Universidades de Latinoamérica donde a menor rango de posiciones se le asocia una mayor puntuación.

Posición	1-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80	81-90	91-100
Puntos	10	9	8	7	6	5	4	3	2	1

**Tabla 5: Asignación de valores según las posiciones de los sitios web.**

Los resultados alcanzados demuestran la veracidad del porqué contar con una buena posición en la Web. Ninguno de los sitios web analizados cuenta con poca presencia de este indicador, lo que evidencia que estos sitios web son bien conocidos, visibles y populares por los internautas. La siguiente gráfica ilustra mejor estos resultados.



**Gráfica 6: Presencia del indicador de posicionamiento en los sitios web latinoamericanos.**

Una vez que se ha podido determinar cuán presentes o no están los indicadores webmétricos definidos en el estudio, el análisis final de todos los resultados de esta primera muestra del estudio, como cabría

esperar, demuestra que el 88.9% de los sitios web de la muestra, han enfocado todo su esfuerzo hacia la aplicación de las técnicas webmétricas necesarias, para alcanzar una adecuada calidad de sus sitios web, que les permita propagarse por todo el Ciberespacio. (Ver Anexo 2)

La mejor manera de lograr esta expansión, volviéndose entonces interesante a los ojos de los internautas de la Web, es presentando sitios web que empleen la utilización de los indicadores webmétricos indispensables para estos fines, o sea, que incluyan un contenido variado, suficiente, actualizado y traducido a varios idiomas. Es mostrando además, sitios web que sean de fácil indexación para los buscadores, posibilitando que se conviertan de esta manera, en los más populares, visibles, accesibles y de mayor impacto en la Web.

### 3.2.2. Análisis de los sitios web de la UCI.

Para el estudio de los sitios web de la UCI se hace uso del WIRE Crawler, el cual es una herramienta estadística que permite observar el comportamiento de una serie de indicadores webmétricos sobre una Web determinada. Aquí no se da un puntaje al comportamiento de los indicadores, simplemente se valora su estado dado una serie de aspectos tales como: la edad y profundidad de sus páginas web, las extensiones encontradas, los enlaces externos e internos, entre otros.

Según el Quinto Estudio Webmétrico realizado en el Departamento Productivo: **Soluciones Informáticas para Internet** (SINI) de la facultad 10, la Web de la UCI contiene aproximadamente 219 sitios, incluyendo los sitios que pertenecen a las tres facultades regionales correspondientes a Artemisa, Ciego de Ávila y Granma. Para el presente estudio se tomó una muestra representativa de 113 sitios, de los cuales se descargaron 164762 páginas. La *Tabla 6* muestra algunos datos interesantes acerca de las páginas web descargadas durante el estudio.

Total de Páginas	164762	
Únicas	156345	94.89 %
Duplicadas	8417	05.11 %
Estática	75414	45.77 %
Dinámica	89348	54.23 %

**Tabla 6: Páginas web descargadas de los sitios web de la UCI.**

A continuación se analiza el comportamiento de los indicadores webmétricos presentes en los sitios web de la UCI y se da una valoración de los mismos.

**Indicador de visibilidad:** (Ver Anexo 3, gráfica 7)

Para medir la visibilidad se tiene en cuenta el grado interno de los sitios web de la colección. Este grado interno no es más que, la cantidad de sitios distintos de la colección que apuntan a un sitio específico desde alguna de sus páginas. En la *Tabla 7* se muestran los sitios web más referenciados.

Dirección del Sitio	Grado Interno (Enlaces Entrantes)
primavera.uci.cu	13
softwarelibre.uci.cu	11
inter-nos.uci.cu	6
web21.uci.cu	6
www.uci.cu	5

**Tabla 7: Sitios web con mayor grado interno de la colección.**

De manera general se podría inferir que un sitio promedio es referenciado aproximadamente desde 1.36 sitios web pertenecientes a la colección analizada. Según los resultados mostrados, se puede apreciar que son muy pocos los sitios con mayor grado interno, y el número que avala este grado interno tampoco es muy representativo. Por lo que a partir de este resultado se deduce que, los sitios web de la UCI en general, no presentan una buena visibilidad web. Esto se debe a la poca cantidad de vínculos que reciben estos sitios.

**Indicador de popularidad y factor de impacto web:** (Ver Anexo 3, gráfica 7)

Para estimar la popularidad y el factor de impacto web de los sitios UCI, se hace uso del grado interno mostrado anteriormente en la *Tabla 7*, ya que este parámetro se encuentra generalmente asociado a la popularidad de un sitio web y contribuye además al impacto web del sitio en sí. Mientras mayor cantidad de enlaces apunten a un sitio web, este se considera popular y existe más probabilidad de que el mismo adquiera un mayor factor de impacto web y por lo tanto sea más conocido y recomendado.

Una vez determinado el grado interno de los sitios web UCI, se deduce que son muy pocos los sitios web que mayor visita adquieren, y una inmensa mayoría recibe muy pocas visitas de usuarios. Esto se debe a que existen distintos intereses en cuanto al abastecimiento de la información en la comunidad UCI. Como

se puede apreciar, los sitios de mayor grado interno son los que presentan información de interés para todos los usuarios de la Universidad, por lo tanto estos son los más visitados.

**Indicador de accesibilidad:** *(Ver Anexo 3, gráfica 8)*

Para el análisis del indicador de accesibilidad se tienen en cuenta: el idioma en el cual se muestra el contenido de las páginas, la profundidad de estas y el grado externo de los sitios web, ya que estos son aspectos relevantes que se deben analizar cuando se quiere mejorar la accesibilidad a la información.

Según la muestra analizada, se obtuvo que los idiomas que más predominan en los sitios web UCI son el español y el inglés, presentes en aproximadamente un 94.96 % y 3.89 % respectivamente. También se encuentran pequeñas muestras de noruego, francés, danés, alemán, portugués, neerlandés, italiano y sueco, todos en conjunto representan un 1.15 % de la colección. Estas pequeñas porciones de muestras van decreciendo en el mismo orden en que fueron mencionadas.

Para la comprensión de los resultados anteriores se puede acotar que, por tener acceso a estos sitios solo usuarios de habla hispana y por ser el español el idioma que más predomina en ellos, mirándolo desde este punto de vista la accesibilidad es significativa. Desde otro ángulo, la Web de la UCI no brinda un escenario óptimo para una población que domine diversos idiomas, además de que no contribuye a una mejor preparación lingüística de los usuarios en cuestión.

Por otro lado, la profundidad lógica de una página web es un indicador que brinda gran importancia al acceso de la información de los sitios web. Este no es más que la cantidad mínima de veces que el usuario tiene que dar clic en un vínculo, para llegar a la misma sin abandonar el sitio web y comenzando desde la portada del mismo. Así mientras menos clic se tenga que dar para acceder a un contenido, más importancia se le dará a este y por lo tanto, menos tiempo invierte el usuario para llegar a la información deseada.

En la colección analizada se obtuvo que unas 641 páginas web se encuentran a solo profundidad 1, representando el 0.39 % del total de páginas; otras correspondiente a 118155 páginas web se encuentran entre una profundidad de 2 y 6, lo cual representa el 71.71 % de la colección. En otro rango de profundidad de 7 a 9, se obtuvo que existen 44989 páginas web, lo que representa un 27.30 % de la



colección. Incluso se pudo observar la existencia de un 0.59 % de páginas correspondiente a 977 páginas web, a una profundidad de 10.

Como los sitios web deben tener el 85 % de sus páginas entre profundidades de 2 y 6 para considerarse con buena accesibilidad al contenido, el resultado obtenido es satisfactorio desde este punto de vista, pues se acerca a dicho porcentaje en estas profundidades. Se pudo apreciar además, que existe un bajo porcentaje de páginas con profundidad mayor de 10, lo cual constituye un valor insignificante a tener en cuenta para este aspecto.

Por tal razón, las páginas relevantes para el usuario no deben encontrarse a grandes profundidades, porque estas podrían nunca ser vistas o encontradas por el mismo. Igualmente, esta tarea puede convertirse en una búsqueda exhaustiva y con pocas comodidades para el acceso a la información, lo cual conlleva a que el usuario abandone el sitio web y salga en búsqueda de resultados concretos y más fáciles de encontrar en otros sitios web. Esto como es lógico, promueve a un descenso de la calidad de aquel sitio web que contiene información valiosa a grandes profundidades de las páginas web.

El grado externo de un sitio no es más que la cantidad de sitios web a los que hace referencia otro sitio específico. Este grado externo permite acceder directamente a una determinada cantidad de sitios web que presentan información de interés para los usuarios, por lo tanto, contribuye a una mayor accesibilidad de la información. Generalmente un sitio de carácter comercial tratará de tener un grado externo pequeño, para evitar que los usuarios abandonen el sitio siguiendo enlaces a otros sitios.

En la *Tabla 8* se muestran los sitios de la colección con mayor grado externo. Esta información ayuda a conocer el nivel de comunicación que existe entre los sitios web de la colección y además contribuye a la visibilidad de aquellos sitios a los que se hace referencia mediante enlaces.

Dirección del Sitio	Grado Externo (Enlaces Salientes)
primavera.uci.cu	13
ucipedia.uci.cu	11
intranet2.uci.cu	11
feu.uci.cu	10
softwarelibre.uci.cu	9
web21.uci.cu	7

**Tabla 8: Sitios web con mayor grado externo de la colección.**

A partir del análisis del grado externo en esta muestra, se obtuvo que un 73.8% de los sitios web que la componen, no apuntan a ningún otro sitio de la colección. Esto no constituye un buen resultado para la accesibilidad de los sitios web de la UCI.

**Indicador de contenido:** *(Ver Anexo 3, gráfica 9)*

El indicador de contenido es un elemento fundamental para la calidad de los sitios web. Aunque existen muchos otros aspectos para medirlo, en este estudio se toman en cuenta la actualidad de las páginas, la cantidad de ficheros ricos y la cantidad de páginas duplicadas y únicas.

La edad de las páginas web que componen un sitio brinda una idea de lo actualizado que puede estar el mismo. De la colección analizada, se obtuvo que un 92.23 % de las páginas web fueron creadas o modificadas en el último año, así mismo un 1.99 % en el año anterior, y solo un 0.02 % que representa a 8 documentos de la colección no han sido actualizadas hace 7 años atrás. El resto de las páginas que representa un 5.76 % han sido modificados en un intervalo de 2 a 6 años atrás. Estos datos muestran un alto crecimiento y dinamismo de las páginas web de la UCI.

Es válido destacar, la importancia de prestar un seguimiento continuo a la edad de las páginas de los sitios web, para mantener siempre una Web lo más actualizada posible, que brinde información detallada y concisa. Este requerimiento aporta mucho al indicador de contenido, pues precisamente de los datos actualizados que muestre un sitio web, el usuario determina su interés en visitarlo o no.

Según el análisis realizado se evidencia que en la Web de la UCI existen diferentes tipos de extensiones, lo cual aporta al indicador de contenido que evalúa, entre otros aspectos, los ficheros ricos en formatos y ficheros media. Se muestra entonces, una representatividad de 566 ficheros de extensiones de video encontradas, sobresaliendo los ficheros .WMV y .MOV con un 56 % y 32.68 % respectivamente. Hay 29408 extensiones de documentos con un 100 % de representatividad de la extensión de fichero XHTML. Se presencian 1241567 ficheros de extensiones de imagen predominando la extensión .GIF con un 68.63 % del total encontrado. Se encontraron además, 197 extensiones de audio predominando los .MP3 con 96.95 % de representatividad.

Todos estos resultados, junto con otras extensiones que fueron encontradas referentes a ficheros comprimidos, extensiones de software, de interfaz de entrada y muchos más, le atribuyen datos importantes al volumen de información de las páginas web. Todas las extensiones aportan a la calidad del contenido, ya que permiten una mejor organización, diseño y presentación de la información de un modo más agradable y eficiente, además de que estos pueden ser reutilizados fácilmente. Los contenidos sin un formato normalizado en la Web, posibilitan acceder a la información de un modo más entretenido y satisfacer los requerimientos y gustos de los más disímiles usuarios.

Por otro lado, se tiene que del total de páginas web descargadas, 8417 páginas se encuentran duplicadas, lo que representa el 5.11 % del total de páginas. Estos datos pueden ser interpretados de distintas maneras, en dependencia de la importancia que adquiere el contenido de la información que se da a conocer. Si la información de las páginas duplicadas contiene un alto valor de conocimiento e interés para los usuarios, entonces esto se toma como un aspecto positivo para los indicadores de contenido y visibilidad, ya que promueven a que el sitio web sea más visitado y conocido. En caso contrario se toma como un aspecto muy negativo, puesto que los usuarios al no satisfacer sus necesidades pierden interés por el sitio, y por lo tanto las probabilidades de alcanzar una alta visibilidad de este se convierten en casi nulas.

Una vez concluido el análisis de los sitios web UCI, se pudo apreciar que estos no presentan en su totalidad un buen comportamiento de los indicadores webmétricos establecidos para la presente investigación. Aunque algunos resultados obtenidos son satisfactorios y promueven a la calidad del sitio, se encuentran ausentes muchos otros parámetros que contribuyen a la total eficacia del mismo.

Un aspecto muy importante, el cual provoca que muchos sitios web de la UCI no presenten buena visibilidad, popularidad e impacto web, además de no cumplir con ciertos parámetros necesarios para cada uno de estos indicadores, es el hecho de no encontrarse posicionados en alguno de los buscadores web por pertenecer a una Web invisible. En este caso, estos sitios solo pueden ser accedidos si el usuario interesado conoce su URL o encuentra algún vínculo hacia este a través de otras páginas web, de lo contrario, como el usuario no puede realizar alguna búsqueda en Internet para encontrar el sitio, es muy poco probable que el mismo sea accedido por los usuarios desde esta vía.

### **3.3. Acciones para mejorar la calidad de los sitios web de la UCI.**

El análisis efectuado en el epígrafe anterior permite sentar las bases para aportar el conocimiento necesario a esta segunda etapa de la estrategia (intervención), de forma tal que permita mejorar la calidad de los sitios web de la UCI.

Luego de detectar qué indicadores webmétricos están presentes o no en los sitios web de la UCI analizados, y cómo influyen estos sobre la calidad de dichos sitios, se elabora un conjunto de acciones que permite erradicar la poca puesta en práctica de los mismos.

#### **1. Acciones para mejorar la presencia del indicador de contenido.**

- Ordenar los contenidos en el sitio web agrupándolos en conjuntos de datos precisos y coherentes, dándoles nombres que lo identifiquen, de forma tal que facilite el sistema de navegación para el usuario durante su búsqueda por las distintas secciones del sitio.
- El contenido de cada página subsiguiente debe guardar relación con su vínculo o link, y ser correlativa con la que le precede.
- Tratar los datos, los detalles, y la información compleja en las páginas subsiguientes y dispuestas de forma lógica.
- Eliminar del contenido todo lo que sea superfluo, pues desvirtúa el mensaje que se quiere dar, por muy genial que este sea.
- Proveer al sitio de información interesante, relevante, precisa y actualizada, que logre satisfacer las necesidades de información del usuario.
- La información del sitio debe mostrar algún grado de validez y confianza, es decir, tratar de confirmar lo que se dice con hechos aceptados para aumentar la credibilidad de la información. Siempre que se pueda, dar a conocer como se han obtenido los datos, para ello se debe indicar la fuente de datos y/o la metodología usada.
- Aportar documentación que apoye a la información dada en el sitio, bien en forma de bibliografía y/o en forma de enlaces a otras páginas, especialmente cuando la información se basa en datos numéricos o estadísticos.
- Lograr que la documentación brindada por el sitio, cumpla los diferentes criterios relacionados con la autoría de la fuente de información (identificación del autor, autoridad, tipo de

organización, afiliación del autor, información del contacto), ya que muchos usuarios le dan bastante importancia a esto para valorar la calidad de la información brindada.

- Presentar calidad en la escritura y sintaxis del contenido mostrado en el sitio: buena gramática y sin errores ortográficos. Pues el hecho de que un texto esté bien escrito es importante para que el proceso de la comunicación se realice de forma clara.
- Tratar que la información en el sitio no se encuentre duplicada, a no ser que sea relevante y de alto valor de conocimiento para los usuarios.
- Si el sitio es comercial o con propósitos financieros, debe dar suficiente información relativa acerca de su historia, productos y estados financieros, para una mayor transparencia respecto a su actividad económica.
- Invitar a los visitantes a realizar comentarios y sugerencias, pues esto da una idea de si el usuario se encuentra satisfecho o no con la información encontrada, y permite realizar futuras transformaciones del contenido en vista de lo que el usuario desea.
- Dejar espacios en blanco entre párrafos para aumentar la legibilidad.
- Si se hace uso de imágenes, se debe valorar bien el objetivo que cumple este, pues su uso puede reforzar o complicar el texto, o incluso no tener sentido y distraer al lector. Las ilustraciones y los gráficos que aparezcan deben añadir valor al sitio, y si se utiliza audio y video deben ser apropiados para el objetivo de este.
- El sitio debe incluir la fecha de creación, y de revisión o de última actualización, para determinar su actualidad.
- Se debe figurar en el recurso, el objetivo y la audiencia a quien va dirigido el sitio. Esto debe realizarse en la página inicial, o proporcionar un enlace desde la página inicial.
- Tener en cuenta a la hora de interpretar y usar la información a publicar, que la misma contenga objetividad, sesgo y equilibrio en su contenido, pues son aspectos que se relacionan con el propósito con el que ha sido escrita la información.
- Determinar el nivel de profundidad y exhaustividad con que se cubre el tema en cuestión, y si la información es completa, qué aspectos se cubren y con qué nivel de detalle. Para ello, se puede comparar el sitio con otros sobre el mismo tema, lo cual puede ayudar a determinar el grado con que la información presentada cubre el tema abordado, y aportar diferentes puntos de vista.

### 2. Acciones para mejorar la presencia del indicador de accesibilidad.

- Siempre que sea posible, lograr una interfaz de usuario multi-plataforma, multi-idioma, y adaptable dinámicamente a los requerimientos de los usuarios, ya que esta sería una vía para terminar con las barreras de accesibilidad.
- Tener versiones de páginas web en otros lenguajes (la mayor cantidad de lenguajes posibles), no solo para las páginas principales, sino para otras secciones importantes del sitio, y principalmente los documentos científicos que son mayormente los más accedidos por los usuarios.
- Comprobar que las páginas web se puedan visualizar de manera correcta en los diferentes navegadores, en caso de que esto no se logre pues se debe realizar un nuevo diseño web que permita la visualización de las páginas en la mayoría de los navegadores.
- El sitio web debe ser accesible para cualquier usuario, con independencia de los programas o del ordenador que este utilice.
- Llevar a cabo un constante mantenimiento de los enlaces a otras páginas o sitios web, con el fin de detectar los vínculos rotos, repararlos o eliminarlos en caso de que estos ya no sean necesarios, de forma tal que se logre la completa accesibilidad hacia otras páginas web.
- Diseñar un sistema de navegación flexible, que ayude al usuario a acceder a la información que necesita en poco tiempo y sin dar muchos clic, ya que si el sitio no es fácil de navegar y de usar, es poco probable que los usuarios lo utilicen. Para ello se pueden incluir herramientas, tales como:
  - ✓ Mapa del sitio, el cual aporta una visión general e incrementa la familiaridad del usuario con respecto al sitio.
  - ✓ Menú permanente del sitio, el cual aporta un rápido acceso a las diversas secciones del sitio.
  - ✓ Buscador por palabras claves, que permitan localizar rápidamente la información interesante del sitio web.
  - ✓ Botones de inicio que conduzcan a la página principal.
  - ✓ Etiquetas visibles e informativas en todas las páginas del sitio web.
  - ✓ Enlaces que envíen al principio y final, o a la página anterior y posterior.

- Hacer uso de un número limitado de gráfico, figuras e imágenes, los cuales lejos de captar usuarios, pueden dificultar la descarga de algunos ficheros y complicar la navegabilidad del sitio.
- La información relevante para el usuario no debe encontrarse en página a grandes profundidades, pues dificulta su accesibilidad en tiempo y corre el riesgo de que nunca sea vista o encontrada por el usuario. Se recomienda que toda información importante se encuentre entre profundidades de 2 y 6.

### **3. Acciones para mejorar la presencia del indicador de visibilidad.**

- Evitar el cambio de nombre de dominio institucional cuando se asigna una URL al sitio, ya que puede generar confusión y tiene un efecto devastador sobre los valores de visibilidad.
- Analizar las estructuras internas y externas de los enlaces en las páginas web, de forma tal que cada enlace funcione correctamente.
- Además de la cantidad de enlaces que apuntan hacia una página web, se debe valorar también la calidad del contenido de esa página para lograr una mayor visibilidad.
- Permitir mayor accesibilidad de los usuarios al sitio en la Red, pues es una forma de lograr que el sitio sea fácilmente visible.
- Conseguir mayor cantidad de enlaces externos desde otros sitios web, que su contenido guarden relación con el sitio en sí.

### **4. Acciones para mejorar la presencia del indicador de popularidad.**

- Elaborar un propio plan de marketing on-line para el sitio, de forma tal que de una imagen de "sitio popular".
- Ofrecer buenos contenidos y servicios para lograr mayor cantidad de visitas al sitio.
- Presentar correctamente sus recursos, actividades y rendimiento global y de esta forma dar a los visitantes una visión real del sitio.
- Lograr que páginas web de otras universidades enlacen a las páginas del sitio, de forma tal que este sea conocido por instituciones prestigiosas y logre mayor popularidad.
- Mantener las páginas web actualizadas pues estas son las más valoradas y visitadas.

### **5. Acciones para mejorar la presencia del indicador de factor impacto web.**

- Lograr que el sitio posea mayor cantidad de enlaces externos, pues de esta forma se obtiene un mayor índice de impacto web.
- Conseguir que la información que brinda el sitio sea de eventual relevancia para el conocimiento de los usuarios, es decir, la información debe ser original, de esta forma, se logrará mayor cantidad de enlaces externos hacia el sitio.
- Atraer el mayor número de usuarios al sitio web de forma tal que se incremente el tráfico en el mismo. Esto provoca que el sitio se haga más popular y adquiera mayor impacto en la Web.
- Familiarizar a los usuarios con el sitio, así como incrementar los servicios que brinda el sitio web.
- Realizar un análisis cualitativo sobre el contenido de los comentarios por parte de los usuarios y, en general, el entusiasmo de la participación del público objetivo. De esta forma se sabrá más sobre la reputación, influencia, compromiso y popularidad del sitio. A partir de los resultados obtenidos se podrá tomar acciones que promuevan a mejorar estos aspectos.

### **6. Acciones para mejorar la presencia del indicador de posicionamiento.**

- Mejorar la calidad del contenido publicado en el sitio web, de forma tal que sea relevante y único, pues esto es de vital importancia para conseguir Rankings altos en los buscadores y evitar que el sitio pueda ser penalizado por contenido duplicado.
- Lograr que el sitio web sea popular entre las comunidades de internautas con intereses específicos.
- Lograr mayor cantidad de links hacia las páginas del sitio, pues de esta forma se tiene más oportunidades de mejorar su posicionamiento, ya que este es uno de los aspectos más valorados en los algoritmos de los buscadores.
- Usar como estrategia para que el sitio posea mayor cantidad de links, el listado de sitios que apuntan a la competencia del sitio que se quiere posicionar. Este puede ser un excelente mapa de aliados, pues si le dan enlaces a la competencia, tal vez estén dispuestos a dárselos al sitio y ya será distancia salvada para lograr una correcta posición.
- Lograr que el código de programación del sitio sea de alta calidad técnica.



- Revisar el código de la página y analizar si la página tiene contenidos en Flash, HTML, Java, JavaScript y otros, de forma tal que se tenga conocimiento de si la página es de fácil acceso para los spider.
- Introducir toda la información de los Meta Tags en las páginas web, pues cada página tiene que tener su Meta Título, Meta Descripción y Meta Keywords. Esto es de vital importancia ya que los buscadores tienen en cuenta todos estos aspectos a la hora de posicionar.
- Lograr que el texto que describe al enlace sea claro, conciso y convincente respecto a lo que se busca, pues los buscadores también evalúan esto a la hora de posicionar, además del número de enlaces que apuntan al sitio y el contenido que les rodea.
- Tratar de posicionar el enlace que apunta al sitio web, lo más arriba posible en la estructura de la página o en el cuerpo del contenido, pues ahí tienen más peso y por lo tanto valen más que aquellos enlaces que son situados al pie de la página.
- Incluir diferentes ítems o criterios de relevancia como son: las palabras claves que aparezcan con la frecuencia adecuada y la ubicación correcta, tipos de links recibidos por el sitio, número de visitantes que recibe la página y contenido.
- La programación básica del sitio web debe contener las palabras claves utilizadas por los usuarios a la hora de hacer una búsqueda sobre un tema específico.

En el presente capítulo se han observado dos estudios webmétricos llevados a cabo en distintas Web con características también diferentes, lo cual ha permitido constatar que nivel de calidad presentan sus sitios web. De esta forma, se crean las bases para venideros estudios webmétricos a nivel nacional e internacional que involucre dicha calidad.

La estrategia metodológica aquí presentada para mejorar la calidad de los sitios web en la UCI, constituye un resultado relevante alcanzado en la investigación; y permite además, guiar a cualquier diseñador desde que inicia el desarrollo de su sitio web, para lograr una óptima calidad de este, permitiéndole ahorrar tiempo y costos en futuras transformaciones como resultado de errores que se pueden encontrar posteriormente.

### Conclusiones Generales

- El tratamiento teórico reflejó que la diversidad de la Web es tan amplia, que los diferentes estudios webmétricos realizados en todo el mundo, no han alcanzado a resolver todos los fenómenos que se encuentran en la misma.
- El conjunto de indicadores webmétricos seleccionados para el análisis de los sitios web lograron ser confiables, precisos y útiles en la elaboración de acciones que contribuirán a la toma de decisiones por parte de los webmáster, para mejorar la calidad de los sitios web en la UCI.
- El diagnóstico realizado en este estudio, demostró que no se aprovechan todas las potencialidades que brinda la UCI para desarrollar, de forma sistemática, sitios web con calidad sobre la base de los indicadores webmétricos. Esto se evidenció a partir de las principales deficiencias detectadas en los sitios web de dicha institución, donde:
  1. El 57.5 % de los sitios web, no presentan una buena visibilidad, popularidad y factor de impacto.
  2. El 31.1 %, de estos sitios web carecen de buena accesibilidad web.
  3. Solo un 3.6% de dichos sitios presentan un contenido desactualizado y duplicado.
- La intervención realizada como parte de la estrategia metodológica propuesta, presenta las condiciones necesarias para que los administradores de los sitios web de la UCI, logren grandes posibilidades de mejorar la calidad de los sitios de esta institución a partir de un conjunto de acciones, contribuyendo a mediano plazo a la solución del problema motivo de esta investigación.

### Recomendaciones

- Realizar en estudios posteriores de la temática la evaluación y la validación de la estrategia metodológica planteada en el presente documento científico, para determinar cómo influye esta en un mejor rendimiento y optimización de los sitios web de la UCI.
- Tomar como muestra de referencia el presente Trabajo de Diploma para futuros estudios relacionados a la calidad de los sitios web, no sólo de la UCI, sino también de otras regiones del país inmersas en el actual desarrollo tecnológico e informatización de la sociedad.
- Incorporar a la estrategia metodológica otros indicadores y acciones relacionadas al objetivo principal, para garantizar la calidad e integridad de los sitios web de la UCI.

### Referencias Bibliográficas

1. **Rodríguez, MsC. Ailín Martínez.** *Indicadores cibernéticos: ¿Nuevas propuestas para medir la información en el entorno digital?* 4, 2006, ACIMED, Vol. 14. [En línea] [Citado en dic de 2009.] [http://bvs.sld.cu/revistas/aci/vol14\\_4\\_06/aci03406.htm](http://bvs.sld.cu/revistas/aci/vol14_4_06/aci03406.htm)
2. **Natalia Arroyo, Pareja y otros.** Cibermetria. Estado de la cuestión. [En línea] Abril de 2005. [Citado el 14 de enero de 2010.] <http://digital.csic.es/bitstream/10261/4296/1/R-17.pdf>
3. **Castaño Muñoz, Wilson.** Cibermetría y medición de la Web: Implicaciones para las bibliotecas. [En línea] 12-15 de mayo de 2009. [Citado en enero de 2010.] <http://www.slideshare.net/colfreepress/cibermetria-wilson-castao>.
4. **Rodríguez Jiménez, Teresa M.** [En línea] [Citado en enero de 2010.] <http://148.202.167.9/cibermetria/wiki/index.php/Portada>
5. **Faba Pérez Cristina, Guerrero Bote Vicente P. y de Moya Anegón Félix.** *Fundamentos y Técnicas Cibernéticas.* España: s.n., 2004. [En línea] [Citado en enero de 2010.] <http://ict.udlap.mx/projects/cudi/datos/tecnicascibermetricas.pdf>
6. **Chávez Sánchez, Guillermo y Estrada Corona, Adrian.** Mati, sobre la letra digital. *Webometría.* [En línea] 30 de Nov de 2009. [Citado en dic de 2009.] [http://www.mati.unam.mx/index.php?option=com\\_content&task=view&id=406&Itemid=49](http://www.mati.unam.mx/index.php?option=com_content&task=view&id=406&Itemid=49)
7. **Rodríguez Aponte, Daysi Y.** Algunos conceptos básicos sobre los indicadores. [En línea] [Citado en dic de 2009.] <http://www.ifrtd.org/spanish/proj/Conceptos%20basicos%20indicadores.doc>
8. Wikipedia. [En línea] [Citado en dic de 2009.] <http://es.wikipedia.org/wiki/Calidad>.
9. **Jiménez, Juan Carlos.** El funcionamiento de los buscadores para los internautas. [En línea] Junio de 2007. [Citado en enero de 2010.] <http://www.internetips.com/articulos/detalle.php?iid=128>
10. RELsoftware. Web Link Validator: The Broken Links Doctor. [En línea] Mayo 2009. [Citado en enero de 2010.] <http://www.relsoftware.com/wlv/>

## REFERENCIAS BIBLIOGRÁFICAS

---

11. ABCdatos. Web Link Validator v5.1 Build 511. [En línea] Mayo 2009. [Citado en enero de 2010.] <http://www.abcdatos.com/webmasters/programa/z3573.html>
12. Giveaway of the Day - Site Content Analyzer. [En línea] nov de 2009. [Citado en enero de 2010.] <http://es.giveawayoftheday.com/site-content-analyzer/>
13. W3C Link Checker. [En línea]. [Citado en enero de 2010.] <http://www.elcodigo.net/cgi-bin/DBread.cgi?tabla=herramientas&campo=0&clave=13&info=1>
14. Softonic. Website Extractor 9.70. [En línea] dic 19, 2007. [Citado en enero de 2010.] <http://website-extractor.softonic.com/>
15. FindMySoft. Teleport Pro 1.60. [En línea] mayo 2009. [Citado en enero de 2010.] <http://teleport-pro.findmysoft.com/>
16. PiredMex. *Webcopier Pro 4.5 & 4.6*. [En línea] dic 2008. [Citado en enero de 2010.] <http://www.piredmex.com/disen-y-creacion-web/webcopier-pro-4-5-4-6-full/>
17. **Gómez, María**. Puro Marketing Diario Digital. España: s.n., [En línea] 2009. [Citado en enero de 2010.] <http://www.puromarketing.com/8/5464/la-escasa-visibilidad-web-caracteriza-sites-principales-cadenas-hoteleras.html>
18. **Vitullo, Nadia Aurora Vanti**. Links Hipertextuais na Comicação Científica. Porto Alegre: s.n., [En línea] 2007. [Citado en enero de 2010.] <http://www.bibliotecadigital.ufrgs.br/da.php?nrb=000607804&loc=2007&l=cd91c7167430503a>
19. **Caraballo, Yeter Pérez; Amoros Aida C. Noda; Herrera, Javier A. Toscazo**. Estudio de la visibilidad de los centros del MES de Cuba, a partir de un análisis webmétrico. Cuba: s.n., [En línea] 2008. [Citado en enero de 2010.] <http://www.elprofesionaldelainformacion.com/contenidos/2008/julio/12.html>
20. **de Ugarte, David**. *El libro del posicionamiento en buscadores*. Madrid; España: s.n., marzo de 2002. [Citado en febrero de 2010.] [http://www.ciberconta.unizar.es/leccion/BUSCAR/el\\_libro\\_del\\_posicionamiento\\_en\\_buscadores.pdf](http://www.ciberconta.unizar.es/leccion/BUSCAR/el_libro_del_posicionamiento_en_buscadores.pdf)

## REFERENCIAS BIBLIOGRÁFICAS

---

21. **Iglesias, M. Sc. Prof. Aux. Ing. Arabel Moráguez.** *¿Cómo seleccionar el tamaño de una muestra para una investigación educativa?* [En línea] abril de 2006. [Citado en dic 2009.] <http://www.monografias.com/trabajos42/seleccion-muestra/seleccion-muestra.shtml>
22. “Sobre la pertinencia de Alexa, sus números y algunas ideas para mejorarlos en su web.” [En línea] Octubre 26, 2008. [Citado en abril de 2010.] [http://www.tecnologiahechapalabra.com/negocios/trafico\\_internet/articulo.asp?i=3147](http://www.tecnologiahechapalabra.com/negocios/trafico_internet/articulo.asp?i=3147)
23. *Garantice Un Aumento Del Tráfico En Su Sitio Web.* [En línea] [Citado en abril de 2010.] <http://goodworkfromhome.com/garantia-aumento-trafico-web.html>
24. **Hernández-Ortega, Blanca, Julio Jiménez-Martínez, M. José Martín-DeHoyosa. Zaragoza.** *Calidad de la información web en la banca electrónica.* España: s.n. [En línea] de 2008. [Citado en enero de 2010.] [http://www.esceap.net/conferences/marketing/2008\\_cp/Materiali/Paper/Fr/HernandezOrtega\\_JimenezMartinez\\_MartinDeHoyos.pdf](http://www.esceap.net/conferences/marketing/2008_cp/Materiali/Paper/Fr/HernandezOrtega_JimenezMartinez_MartinDeHoyos.pdf)
25. **Hernández, Y. Mondelo, Madrugá, Yuley Días.** “*Características de la Web de la Universidad de las Ciencias Informáticas*”. Ciudad Habana: mayo de 2009. [Páginas citadas 74-82.]

## Bibliografía

**Codina, Lluís.** *El posicionamiento web y su importancia para las empresas.* marzo, 2007. [En línea] [Citado en marzo 2010.] <http://www.lluiscodina.com/posicionamiento2007v2.pdf>

**Codina, Lluís.** *Evaluación de recursos digitales en línea: conceptos, indicadores y métodos.* 1, España : s.n., 2000, Vol. 23. [En línea] [Citado en febrero de 2010.] <http://www.lluiscodina.com/metodos/eval2000.doc>

**Codina, Lluís.** Posicionamiento Web: Conceptos y Ciclo de Vida. "Hipertext.net", núm. 2, 2004. [En línea] [Citado en octubre de 2009.] <http://www.hipertext.net/web/pag216.htm>

**Colectivo de autores:** Metodología de la investigación social. Segunda Edición.

Definiciones de Robots de Búsqueda. [En Línea] [Citado en enero de 2010.] <http://www.abcdatos.com/buscadores/robot.html>

**Faba Pérez, Cristina, Guerrero Bote, Vicente P y de Moya Anegón, Félix.** *Fundamentos y Técnicas Cibernéticas.* España : s.n., 2004. [En línea] [Citado en enero de 2010.] <http://ict.udlap.mx/projects/cudi/datos/tecnicascibermetricas.pdf>

Hispanmedios. Page Rank. [En línea] 9 de Diciembre de 2006. [Citado en noviembre de 2009.] <http://www.hispamedios.com/articles/id27-pagerank>.

**Ibarra Martín, Francisco:** Metodología de la investigación. Editorial Félix Varela, La Habana, 1999. [Citado en abril de 2010.]

**Aguillo, Isidro F.** Cibermetría Introducción teórico-práctica a una disciplina emergente. [En línea] Febrero de 2003. [Citado en Marzo de 2010.] <http://internetlab.cindoc.csic.es/cursos/cibermetria.pdf>

**Jiménez, Juan Carlos.** Una buena posición en los buscadores. [En línea] Abril de 2007. [Citado en enero de 2010.] <http://www.internetips.com/articulos/detalle.php?iid=125>

**Barrera, Lic. Carlos Robert, MsC. Sonia Núñez Amaro y Lic. Daniel Motola Pedroso.** *Evaluación de sitios Web en Internet. Propuestas para la evaluación de sitios web de bibliotecas públicas y de salud .* 4,

2006, ACIMED, Vol. 14. La Habana, Cuba. [En línea] [Citado en marzo 2010.]  
[http://www.bvs.sld.cu/revistas/aci/vol14\\_4\\_06/aci04406.htm](http://www.bvs.sld.cu/revistas/aci/vol14_4_06/aci04406.htm)

**Pérez Pareja, Víctor Manuel y Natalia Arroyo Vázquez.** *Metodología para la obtención de datos con fines cibernéticos.* 2002. [En línea] [Citado en enero de 2010.]  
<http://internetlab.cindoc.csic.es/varios/Methodolog%EDa%20datos%20cibern%EA9tricos.pdf>

**Grau, Ricardo Abalo, Cecilia Correa Valdés y Mauricio Rojas Betancur.** *Metodología de la Investigación.* Ibagué : s.n., 1999. [Citado en abril de 2010.]

**Rodríguez Gómez, Gregorio y otros:** *Metodología de la investigación cualitativa.* Santiago de Cuba, febrero 2002. [Citado en abril de 2010.]

**Ugarte, David de.** *El libro del posicionamiento en buscadores.* Madrid; España : s.n., marzo de 2002. [En línea] [Citado en febrero de 2010.]  
[http://www.ciberconta.unizar.es/leccion/BUSCAR/el\\_libro\\_del\\_posicionamiento\\_en\\_buscadores.pdf](http://www.ciberconta.unizar.es/leccion/BUSCAR/el_libro_del_posicionamiento_en_buscadores.pdf)