

Depósito legal: ppi 201502ZU4635

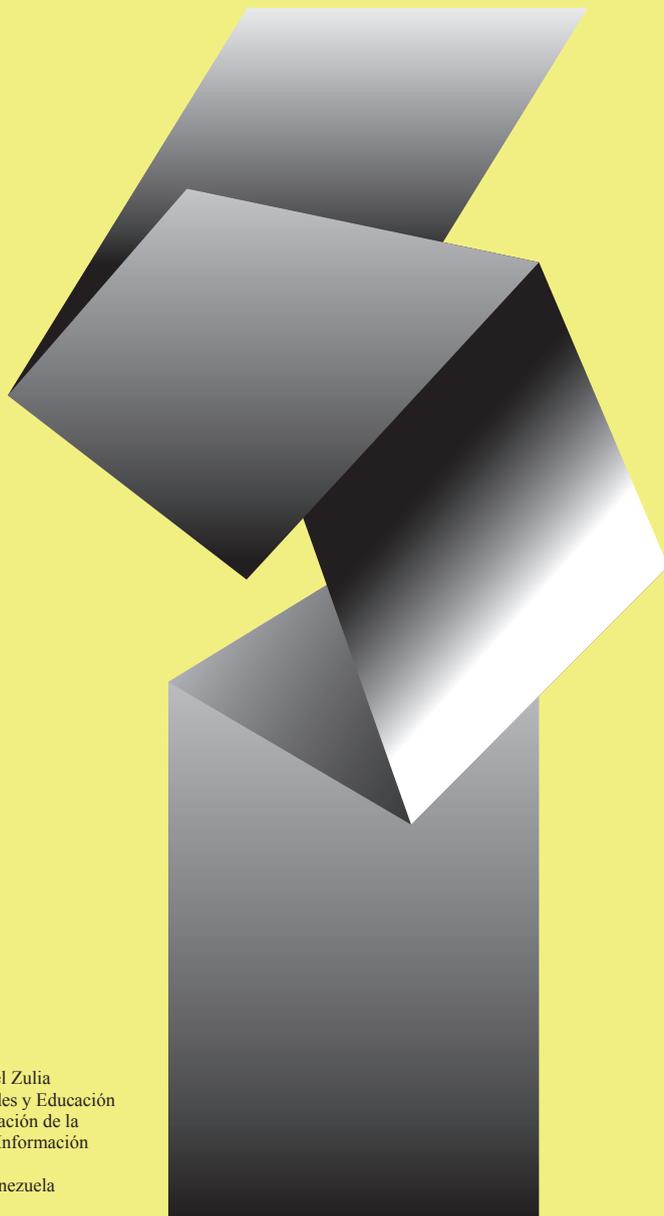
Esta publicación científica en formato digital es continuidad de la revista impresa

Depósito Legal: pp 200402ZU1627 ISSN:1690-7582

# Q U Ó R U M

## ACADÉMICO

Revista especializada en temas de la Comunicación y la Información



Universidad del Zulia  
Facultad de Humanidades y Educación  
Centro de Investigación de la  
Comunicación y la Información  
(CICI)  
Maracaibo - Venezuela



## Hacia las técnicas no invasivas en comunicación

*Ángel Emiro Páez Moreno<sup>1</sup>*

### Resumen

El objetivo del artículo es discutir la viabilidad de las denominadas técnicas no invasivas (Unobstrusive Measures) en el campo de la comunicación. Las técnicas no invasivas sirven como complemento de las técnicas de obtención directa, son adaptable en situaciones difíciles o peligrosas y promueven metodologías creativas. Se trabaja con datos encontrados, datos capturados y datos recuperados (registros de servicio, registros personales, episodios). Algunas de sus ventajas en el ciberespacio son: muchas fuentes, posibilidad de búsqueda, los datos ya están digitalizado y listos para procesar, se acaba el problema de la representatividad de los datos demográficos de los usuarios y son menos limitada geográficamente. El análisis de contenido tiene tres momentos: Primer momento: Europa Central y Gran Bretaña (Finales del S. XIX y principios del S. XX). Investigaciones centradas en la efectividad de la prensa y el cine (Henry de Noussance y P. Stoklossa). Segundo momento: Segunda Guerra Mundial. Tesis de los efectos poderosos. Se consolida el análisis de contenido como técnica usual en los trabajos de investigación. Tercer momento: 1960 en adelante (uso de computadoras y ampliación de comunicación no verbal). Se propone la cibergrafía (técnica para medir a los medios de comunicación social cibernéticos) como técnica de investigación no invasiva.

**Palabras clave:** investigación, comunicación, comunicología, técnicas de investigación.

---

Recibido: Mayo 2018 - Aceptado: Agosto 2018

<sup>1</sup> Profesor titular de La Universidad del Zulia (Venezuela). Director del Centro de Investigación de la Comunicación y la Información. Investigador del Programa de Comunicación Social de la Universidad de Boyacá (Colombia). [angel.paez@hdes.luz.edu.ve](mailto:angel.paez@hdes.luz.edu.ve) – [aepaez@uniboyaca.edu.co](mailto:aepaez@uniboyaca.edu.co)



## *Towards non-invasive communication techniques*

### **Abstract**

The objective of the paper is to discuss the viability of the so-called non-invasive techniques (Unobstrusive Measures) in the field of communication. Non-invasive techniques serve as a complement to direct obtaining techniques, are adaptable in difficult or dangerous situations and promote creative methodologies. It works with data found, data captured and data recovered (service records, personal records, episodes). Some of its advantages in cyberspace are: many sources, the possibility of searching, the data is already digitized and ready to process, the problem of the representativeness of the demographic data of the users is finished and they are less geographically limited. Content analysis has three moments: First moment: Central Europe and Great Britain (Late 19th Century and early 20th Century). Investigations focused on the effectiveness of the press and the cinema (Henry de Noussance and P. Stoklossa). Second moment: Second World War. Thesis of powerful effects. Content analysis is consolidated as a usual technique in research work. Third moment: 1960 onwards (use of computers and expansion of non-verbal communication). Cybergraphy is proposed (technique to measure cybernetic social media) as a non-invasive research technique.

**Keywords:** research, communication, comunicología, investigation techniques.

### **Introducción**

Éste trabajo se produce en el contexto de la tradición de investigación en comunicación del Departamento de Investigación de la Escuela de Comunicación Social de la Universidad del Zulia, y muy especialmente en la producción intelectual de sus profesores en el tema de los métodos y técnicas de investigación aplicadas al campo de la comunicación, preocupación siempre válida entre quienes hemos tenido que formarnos con obras provenientes de otras ciencias sociales (especialmente la psicología y la sociología).

A modo de sinopsis cronológica, y de introducción a éste trabajo que se ocupa de la medición de contenidos en el ciberespacio, señalamos

algunos de los artículos y obras que evidencian la producción del grupo de investigadores arriba señalado:

- Cely Álvarez, Adriana. (2004). Cibergrafía: Propuesta teórico metodológica para el estudio de los medios de comunicación social cibernéticos. *Opción*, abril, 101-118.
- Páez, Ángel (2004). Una aproximación cibergráfica al periodismo venezolano en la red. Resúmenes II jornadas de investigación del CICI. , p.18 – 19, Maracaibo - Venezuela.
- Cely, Adriana (2006). Estructura editorial de los cibermedios en Venezuela. En Rojano, Miladys Coord. (2006), Diez años de periodismo digital en Venezuela, pp. 119-148.
- Méndez, Ana Irene (2007). Métodos y técnicas de investigación aplicadas a la comunicación. Maracaibo: Ediciones del Vice Rectorado Académico de La Universidad del Zulia.
- Prieto, Leticia (2009). Métodos cuantitativos aplicados a la comunicación social. Material en proceso de publicación. Universidad del Zulia.

Sólo haremos referencia a las técnicas de medición de contenidos, debido a la importancia que solemos darle los investigadores de la comunicación al mensaje, en lo cual, como veremos, se produce un cambio radical de actores (emisores-receptores) entre la investigación hemerográfica y los análisis cibergráficos.

### **Hemerografía, audiografía e iconografía: las técnicas para medir a los medios tradicionales (prensa, radio, TV y Cine)**

Los primeros análisis de prensa los observamos a finales del Siglos XIX y principios del Siglo XX en Europa Central y Gran Bretaña, cuyo interés de centraba en medir la efectividad de la prensa y el cine. Posteriormente, en un segundo momento, durante la Segunda Guerra Mundial, ya habiendo demostrado la efectividad de los medios, se desarrolla la tesis de los efectos poderosos, consolidándose el análisis de contenido como técnica usual en los trabajos de investigación. En un tercer momento (hasta la actualidad), la

aparición del microcomputador en los años 60's, permite la popularización y automatización de la medición mediada por ordenadores.

El objetivo del análisis de contenido es medir cuantitativa y cualitativamente la información relacionada con el tema con un tema específico, un evento político, por ejemplo, cubierto por los diarios durante un período determinado por el diseño muestral. En la medición hemerográfica que ejemplificamos se atiende principalmente a la morfología del periódico, que puede complementarse con el análisis de contenido según las necesidades del investigador (Méndez, 2015).

En la medición de una pieza, que constituye la unidad de medición, cada noticia, artículo de opinión o aviso constituye un "caso". Si, por ejemplo, se quiere determinar cuánto espacio informativo, publicitario y de opinión recibió un candidato en una elección se medirán las piezas que hagan referencia a ese candidato. Si una noticia tiene llamada en primera página, se medirá cada parte como un caso separado, esto es, el llamado y la noticia en sí se considerarán como casos separados.

De acuerdo a Méndez (2015), en la medición hemerográfica se consideran los siguientes aspectos o categorías (tabla 1):

**Cuadro 1**  
**Categorías hemerográficas**

<b>Categorías</b>	<b>Descripción</b>
Mancha	La mancha o espacio ocupado por una información/artículo/aviso se mide utilizando como unidad el centímetro cuadrado (cm <sup>2</sup> ). Cuando los periódicos eran compuestos manualmente en columnas de ancho fijo, la medición se realizaba utilizando una unidad especial denominada centímetro/columna (cm/col), pero la composición digital de los periódicos ha dejado atrás el diseño basado en las "ocho columnas". La mancha se calcula multiplicando el ancho que ocupa la pieza en centímetros por su altura en centímetros. Así, una pieza que mida 5 cms. de alto por 12 de ancho tiene una mancha de 60 cms <sup>2</sup> .

Mancha de página	<p>Es el espacio en centímetros cuadrados resultante de multiplicar el ancho por el alto de la página. En un periódico estándar cuya medida vertical sea 54 cms. y el ancho 30 cms. la mancha de página será:</p> $\text{Mancha página} = 30 \text{ cms} \times 54 \text{ cms.} = 1.620 \text{ cms}^2$ <p>El cálculo de la mancha de página puede utilizarse, por ejemplo, para determinar comparativamente el espacio proporcional dedicado a información y el ocupado por la publicidad.</p>
Mancha o superficie total	<p>El espacio total o mancha total del ejemplar se obtiene multiplicando la mancha de página por el número de páginas del ejemplar. Así, siguiendo el ejemplo anterior, en un ejemplar de 36 páginas la mancha total será:</p> $\text{Mancha total} = 1.620 \text{ cms}^2 \times 36 \text{ páginas} = 58.320 \text{ cms}^2.$ <p>La medida de la mancha total es útil en estudios donde, por ejemplo, se desea comparar la distribución porcentual del espacio del periódico entre el espacio informativo, el de opinión y el de publicidad.</p>
Mancha o superficie de redacción	<p>Ésta es igual a la superficie o mancha total menos la mancha publicitaria.</p>
Elementos morfológicos	<p>Hay tres grandes categorías de elementos morfológicos a considerar: el título, las ilustraciones y los textos propiamente dichos.</p> <p>El título comprende: Antetítulo</p> <p style="text-align: center;">Título</p> <p style="text-align: center;">Subtítulo</p> <p style="text-align: center;">Intertítulo</p> <p>Las ilustraciones que pueden ser a color o blanco y negro, comprenden fotografías, dibujos, diagramas, cuadros estadísticos, gráficos, infografías y caricaturas.</p>

	<p>El texto propiamente dicho o cuerpo de la noticia.</p> <p>En los estudios hemerográficos se puede partir de una categorización general de temas relacionados bien con géneros periodísticos o con las secciones en que se agrupan las noticias en el periódico bajo estudio. Se pueden hacer estudios en que se relaciona la superficie dedicada a un evento o tema con la superficie total, con la superficie publicitaria o, en otros casos al interior de secciones de la superficie redaccional.</p>
Fuente	<p>Es el actor (persona, institución, partido) que da origen a la información. Si se trata de una pieza publicitaria, la fuente será la persona, empresa, institución o partido que endorsa la publicidad. En el caso del género de opinión, la fuente será el articulista. Si no aparece expresamente identificada, se registrará la fuente como “no identificada”. Aquí el periodista, a menos que sea el autor de una columna de opinión, no se considera fuente. El periodista recoge y procesa la información suministrada por la(s) fuente(s). Es recomendable registrar en una lista aparte las fuentes según vayan apareciendo y se las irá codificando sucesivamente cuidando de que no aparezca la misma fuente con dos numerales (o códigos) diferentes ni dos fuentes con el mismo numeral. En la ficha se registrará el código correspondiente a la(s) fuente(s) de su pieza específica. En la medida de lo posible y según las necesidades de la investigación, se tratará de crear categorías amplias que incluyan fuentes afines. Una fuente puede hablar de sí misma. En este caso, la fuente y el objeto son los mismos.</p>
Partido/ Organización	<p>En el caso de temas políticos es la variable que indica la filiación política o adscripción a una organización de la fuente que suministra la información o la organización política o el organismo que avala un texto publicitario. Si el texto no indica expresamente el partido/organización al que se adscribe la fuente, se registrará la pieza bajo la categoría ‘no identificado’. Si no aparece la categoría del partido o grupo en los valores de la variable, se registrará bajo ‘otro’</p>

Objeto del mensaje	<p>El objeto del mensaje es de quién o de qué se habla en el texto. En una medición, el objeto del mensaje es el actor (político, gremial, vecinal, gubernamental, religioso, ciudadano, académico) al cual alude el texto. Si el texto del mensaje se refiere a más de un actor, entonces se registrará la información según sea, al objeto1, objeto2, objeto 3. Si el texto señala no a un actor sino a un evento éste será el objeto. También aquí es aconsejable registrar en lista aparte los objetos según vayan apareciendo, con la advertencia que se señaló para la fuente de no repetir numerales para dos objetos diferentes ni asignar numerales diferentes a un mismo objeto. También aquí es recomendable crear categorías amplias como se señaló para las fuentes.</p>
Tratamiento	<p>El tratamiento se deduce del contenido del mensaje y constituye por eso mismo un tipo de análisis cualitativo. Si el texto se refiere denotativa o connotativamente mediante el uso de formas verbales, adjetivales o contextuales de manera positiva al objeto del mensaje, se registrará un tratamiento positivo. Si, por el contrario, la referencia denota o connota valores negativos, se registrará el tratamiento como negativo. Cuando no haya calificaciones positivas ni negativas, entonces se registrará como neutral.</p> <p>Si hay más de un objeto del mensaje, el tratamiento se referirá a cada uno de ellos: al objeto1 concuerda con el tratamiento1; el objeto2 se relacionará con el tratamiento2; y el objeto3 se corresponde con el tratamiento 3 (ver modelo de ficha). En otras palabras los valores de la variable “objeto” deben correlacionarse con los respectivos valores de la variable “tratamiento del objeto”</p>

Fuente: Méndez (2015)

En el caso de la medición en medios audiovisuales (iconografía), Méndez (2015) plantea que independientemente de la naturaleza y propósitos de la medición, en todos los casos se recomienda grabar el (los) programa(s) y completar la ficha sólo después de haber hecho un análisis del contenido de, al menos, un programa para identificar los posibles elementos a incluir en las variables y especialmente en los elementos valorativos.

Los elementos valorativos propuestos en el modelo se ordenan de manera descendente para indicar el peso que cada uno de ellos tiene en la importancia que la televisora atribuye a una noticia. Cuando un evento considerado importante ocurre durante el tiempo en que sale al aire el noticiero, la televisora lo cubre en vivo mediante su equipo móvil de microondas o por satélite. De allí que en la jerarquización asignemos el mayor valor a la transmisión en vivo. La transmisión de la noticia por telefonía celular la jerarquizamos en segundo lugar, pues el hecho de que el departamento de producción envíe a un periodista a cubrir el evento en tiempo real indica que le concede importancia (Méndez, 2015).

La medición de un noticiero radial (audiografía) guarda similitudes con la del noticiero televisivo, excepto por lo referido a los elementos de imagen. Las secciones del noticiero varían no sólo con respecto a las de la televisión sino de una emisora a otra, o de una cadena radial a otra. Quien haga una medición de noticiero radial en nuestro medio, aun no siendo residente en la ciudad, pronto caerá en cuenta de que la mayoría de las emisoras no tienen departamento de producción de noticieros y se limitan a leer la prensa escrita del día. Son las cadenas radiales las que cuentan con recursos para mantener un personal de planta y que tienen la organización y los equipos necesarios para recoger la información directamente de las fuentes. Estas circunstancias deben tomarse en consideración al diseñar una ficha de radio. Pero esto puede hacerlo un investigador experimentado. Si la intención es con fines didácticos se hace necesario entrenar a los estudiantes para que aprendan a distinguir entre las noticias leídas y las noticias producidas. De allí que es recomendable que cuando vayan a hacer una medición de radio, lean la prensa del día de manera que tengan la referencia necesaria para hacer la distinción (Méndez, 2015).

### **La cibergrafía como técnica de investigación no invasiva**

Las técnicas de investigación no invasivas o no obstrusivas, no pretenden excluir a las clásicas técnicas de las ciencias sociales como las entrevistas y los cuestionarios; se considera que son complementarias en la medida que permiten recolectar datos inusuales como la basura, graffiti y obituarios, así como otros más convencionales, como las estadísticas publicadas;

también permiten fortalecer hipótesis e investigar temas delicados como el comportamiento sexual (Kalman, s/f).

La cibergrafía se clasifica como una técnica no invasiva, en tanto permite estudiar a los medios digitales sin tener que recurrir a los usuarios. Cely y Neüman (2000) realizaron una propuesta teórico-metodológica para analizar a los medios de comunicación social cibernéticos (o cibermedios). Este gran paso representa una verdadera innovación en la investigación de las ciencias de la comunicación, ya que hasta el momento solo existían metodologías inadecuadas para la plataforma digital como la hemerografía y la audiografía. Al respecto, Cely (2004) explica que:

"La cibergrafía, tal como lo determina el prefijo ciber, devenido de cibernética, estaría conformada por representaciones simbólicas del mundo físico, que sólo existen en el formato electrónico y no en el mundo de los átomos, y, transitan en el espacio virtual creado por las redes de computadoras y líneas de comunicaciones; el sufijo grafía está referida a la descripción gráfica de las representaciones cibernéticas. Dentro de los elementos definitorios de la interactividad y sus niveles, la virtualidad y el carácter hipermediático, la estructura de los MCSC (Medios de Comunicación Social Cibernéticos) comprende a su vez características morfológicas, de contenido, actualización, recursos humanos, recursos publicitarios, finanzas, elementos de Interactividad, elementos de documentación y personalización".

Cely (2004) propone las siguientes categorías cibergráficas para estudiar los medios digitales (tabla 2):

**Cuadro 2**  
**Categorías cibergráficas**

Categoría	Descripción
Morfología	El estudio de la morfología está referido a indagar sobre la estructura técnica que permite la consulta de las páginas de los MCSC y, a la presentación gráfica de la primera pantalla; "el estudio de estos elementos permite comprender cuál es la estructura predominante...y si ésta se aleja o se acerca de la forma de la fuente original" (Cely, 1999:116).

	<p>de esta manera tenemos una morfología de registro, y otra morfología de presentación. En la primera se establece un estudio de la información que describe las características técnicas de donde procede la información y esto comprende: ubicación del servidor, servidores imágenes, visitas registradas, procedencia de los visitantes, fecha de creación, tamaño del sitio, cantidad de enlaces internos y externos, formas de acceso (gratis o registro prepagado).</p> <p>Se trata de recabar la mayor cantidad de datos que describan el sitio y la forma como se almacena la información, esto comprendería datos de registro de la empresa y constitución jurídica, sistemas de distribución y alianzas. Es importante conocer las megafusiones y complejas uniones que establecen las grandes empresas de medios con servidores de Internet, que sin duda determinan el norte de ese MCSC y, sobre todo, los niveles de credibilidad, ya que en un mundo donde cada vez más abundan las fuentes de información, sobrevivirá el que goce de mayor veracidad.</p> <p>El aspecto morfológico que comprende el estudio de la presentación de la primera página busca saber cómo se distribuye la información en la primera pantalla: los recursos iconográficos utilizados, distribución del contenido, elementos de continuidad con las páginas internas, manejo de secciones o subsites.</p>
Contenido	<p>En esta categoría se estudia el tipo de información presente en los MCSC, tales como: áreas temáticas, tipo de publicaciones según la información (de corte generalista o especializadas en algún tema). Comprende a su vez servicios al usuario; es decir, elementos de valor agregado, concursos, información útil que Pérez y Peréa (1998:3) denominan elementos de funcionalidad; incluye también información local, cálculos de inversiones de la bolsa, resultados deportivos, entretenimiento (servicios de correo gratuito, concursos, entre otros).</p> <p>En los MCSC los contenidos están definidos por el usuario; es decir, ya no se está ante el sistema broadcasting (modo tradicional de transmisión de la información), sino con el sistema narrowcasting, que está definido por los destinatarios finales y no por el emisor.</p>

	<p>Esto obliga al periodista a una hipersegmentación del mensaje, no de forma individualizada, sino según un perfil de intereses propuestos por el usuario, y el medio debe entonces ofrecer contenidos que satisfagan esas necesidades.</p>
Actualización	<p>Se refiere a la periodicidad con la cual el MCSC presenta nuevas informaciones. En este medio de redes digitales es posible añadir información de última hora o presentar datos en tiempo real, sin interrupciones, sino como información complementaria, por eso esta dimensión se trabaja con la renovación o edición parcial o total de la información noticiosa, tipo de actualización y número de veces que actualiza.</p>
Interactividad	<p>Está referida a cómo se manejan los niveles de interactividad. Si el sitio del MCSC busca niveles de consulta interactivo o un nivel interactivo de conversación (Jensen, 1998), eso se logra analizando la presencia de elementos para una respuesta en tiempo diferido como se logra a través del correo electrónico o una respuesta inmediata, tiempo real, como son salas de conferencia o chat room (6). De igual forma, si el sitio invita a participar con productos generados por el usuario, encuestas, o coloca el correo de sus redactores para establecer algún contacto con sus productores.</p>
Elementos de documentación	<p>En esta dimensión se busca comprobar si el medio realmente está haciendo uso de una estructura hipermediática que permita recrear la información. Tal vez una de las potencialidades de este medio es la posibilidad de enriquecer la noticia con documentos de extensión recreativos del problema sin costo mayor; de tal forma, se busca detectar si existen bases de datos que pueden ser consultadas a través de archivos electrónicos, directorios de búsqueda, enlaces que amplíen la información en voz, imagen o texto, formato de avance de noticias, servicios de voz, videos, niveles de lectura según la cantidad de vínculos para ampliar el contenido informativo y enlaces hipertextuales que permitan contextualizar la noticia con sitios especializados.</p> <p>La documentación nos permite encontrar las posibilidades de almacenamiento y consulta de información, con lo cual puede ponerse al alcance del usuario información contextualizada, además de profundizar en temas de interés, así como ofrecer bases de datos del mismo medio de comunicación” (Cely, 1999:138)</p>

	<p>Es la posibilidad del medio de responder a gustos e intereses de los usuarios a través de páginas de registro: en el sistema Pull el usuario consulta el sitio tomando la información de su gusto. En el sistema Push: el usuario selecciona unas categorías previas que le son enviadas posteriormente a su buzón de correo o se elabora una página personalizada, si posee versión de la publicación para ser consultada fuera de línea, servicio de titulares al buzón de correo, suscripción a secciones específicas. Entre más personalizado y documentado esté el MCSC mayor es su nivel de innovación, ya que esta última está referida a la posibilidad de que se introduzcan nuevos productos, mejoras o nuevos procesos para organizar y abrir nuevos mercados e irrumpir en el proceso de crecimiento económico.</p>
Usabilidad	<p>Es definida por la norma ISO-9241:11 de 1993 como “la facilidad de uso de una aplicación informática” (Gutiérrez y Páez, 2012: 32). La palabra, que deriva del inglés usability, tiene como objeto definir el grado de facilidad o dificultad que genera el diseño del invento tecnológico a usar y analizar, así como del tipo de proceso educativo que su usuario deba adquirir para usar dicha tecnología. En el caso de la plataforma web Gutiérrez (2009: 47) explica que “la usabilidad permite que los usuarios puedan realizar las tareas a través del entorno gráfico de la interfaz del sitio, es decir, la usabilidad será responsable de que el usuario cumpla su cometido o no”.</p>
Accesibilidad	<p>Característica íntimamente ligada con la usabilidad. Se refiere a la cantidad de usuarios que pueden utilizar correctamente el medio, “la accesibilidad web significa que personas con algún tipo de discapacidad van a poder hacer uso de la web.</p> <p>En concreto, al hablar de accesibilidadWeb se está haciendo referencia a un diseño web que va a permitir que estas personas puedan percibir, entender, navegar e interactuar con la web, aportando a su vez contenidos” (Lawton, 2005:1). A partir de estas características definidas en la cibergrafía, Páez (2009) agrega otras variables que se presentan extensivamente en los blogs, y que pueden ser aprovechados para ser investigados dentro de su propuesta metodológica: Sindicación, que es la presencia del blog dentro de algún catálogo y los canales RSS que posee; la Publicidad contextual, que son sistemas publicitarios que ofrecen</p>

	empresas como Google; y el uso que hace el blog a estudiar de su sistema de gestión de contenidos, tanto desde el punto de vista tecnológico (habilitar comentarios moderados, uso de mensajerías instantáneas, uso de etiquetas), como desde el punto de vista de su creador (tipo de contenidos y elementos dentro de estos).
--	---

Fuente: Cely (2004)

## **A manera conclusión: hacia la minería de datos como técnica de investigación**

Las técnicas de investigación en comunicación deben responder a los cambios que experimentan los medios de comunicación. Sin la cibergrafía, es prácticamente imposible estudiar a los cibermedios o cualquier otro medio digital como los sitios de gobierno electrónico. Con la migración de los lectores-usuarios desde los medios tradicionales hasta las redes sociales, se hace necesario repensar y ensayar nuevas técnicas de investigación como la minería de datos. La minería de datos nació con la idea de aprovechar dos cosas: la ingente cantidad de datos que se almacenaban en áreas como el comercio, la banca o la sanidad, y la potencia de los nuevos ordenadores para realizar operaciones de análisis sobre esos datos. El data mining permite encontrar información escondida en los datos que no siempre resulta aparente, ya que, dado el gigantesco volumen de datos existentes, gran parte de ese volumen nunca será analizado. La minería de datos es un proceso de identificación de información relevante extraída de grandes volúmenes de datos, con el objetivo de descubrir patrones y tendencias estructurando la información obtenida de un modo comprensible para su posterior utilización (Mata, 2017; citado por Vallejo, Guevara y Medina, 2018).

Vallejo, Guevara y Medina (2018) explican que la minería de datos hace referencia a un conjunto de algoritmos que permiten la identificación de patrones útiles y novedosos que se encuentran “ocultos” en grandes bases de datos. Previo al desarrollo de los modelos es necesario surtir 4 pasos:

- a) Entendimiento del problema: se define la pregunta que se quiere resolver y se establecen los objetivos específicos a trabajar.

- b) Selección de datos: se determina qué bases de datos pueden ser utilizadas para analizar el problema y se consolidan los conjuntos de información.
- c) Pre-procesamiento y limpieza: se busca manejar los datos faltantes junto con los valores atípicos que se identifiquen en las bases de datos.
- d) Transformación de datos: se busca generar nuevos indicadores que puedan ser más representativos del fenómeno analizado o contener mejor información que la de las variables disponibles por sí mismas, tomando la información acumulada hasta este momento. (Basagoiti, 2007)

## Referencias

- Basagoiti Astigarraga, R. (2007). **Técnicas de minería de datos aplicadas a series temporales bursátiles**. Madrid: Mondragón Unibertsitatea.
- Basagoiti Astigarraga, R. (2007). **Técnicas de minería de datos aplicadas a series temporales bursátiles**. Madrid: Mondragón Unibertsitatea.
- Cely, A. (1999). La Innovación Tecnológica en las Publicaciones Electrónicas Diarias de Venezuela. Trabajo especial de grado para optar al título de Magister en Ciencias de la Comunicación. Universidad del Zulia.
- Cely, A. (2004). Cibergrafía: Propuesta teórico metodológico para el estudio de los medios de comunicación social cibernéticos. **Opción**, abril, 101-118.
- Cely, A. y Neüman, M. (2000). Caracterización de los nuevos medios cibernéticos de comunicación social en Venezuela. **Revista Comunicación**, 111, Centro Gumilla, Caracas.
- Gutiérrez, Gustavo (2009). Usabilidad de las páginas iniciales de cibermedios venezolanos. Trabajo especial de grado presentado para optar al Título de Magister Scientiarum en Ciencias de la Comunicación, mención Nuevas Tecnologías de la Información. Universidad del Zulia.

- Gutiérrez, Gustavo y Páez, Ángel (2012). Usabilidad en cibermedios nativos e inmigrantes. **Revista Temas de Comunicación** # 24, disponible: <http://revistasenlinea.saber.ucab.edu.ve/temas/index.php/temas/article/view/736> (Consulta: 2013, mayo 18)
- Henry, F.; Vallejo, B.; Edelmira, G.; Segundo, R. (2018). Minería de Datos. **Revista Científica Mundo de la Investigación y el Conocimiento**. Vol. 2 núm., especial, mayo, ISSN: 2588-073X, 2018, pp. 339-349
- Jensen (1998). Interactivity. Nordicom Review, **Nordic research on media and communication review**, No. 19 (2)
- Kalman, Y. (s/f). Unobtrusive Methods for Social Science Research. Consultado el 03-10-18. Obtenido de: <https://slideplayer.com/slide/6090651/>
- Lawton, Shawn (2005). WAI: estrategias, pautas, recursos para hacer la Web accesible a personas con discapacidad, disponible en: <http://www.w3c.es/Traducciones/es/WAI/intro/accessibility> (Consulta: 2013, mayo 18)
- Mata, R. (13 de Enero de 2017). icemd.com. Obtenido de icemd.com: <https://www.icemd.com/digital-knowledge/articulos/mineria-datos-proceso-areas-se-puede-aplica/>
- Méndez, A. (2015). **Metodologías y técnicas de investigación aplicadas a la comunicación**. Maracaibo: Ediciones del Vicerrectorado Académico de la Universidad del Zulia.
- Páez, Ángel (2004). Una aproximación cibergráfica al periodismo venezolano en la red. **Resúmenes II jornadas de investigación del CICI** , p.18 – 19, Maracaibo - Venezuela.
- Páez, Ángel (2009). De la hemerografía a la blogografía: Evolución de las técnicas de medición de los medios de comunicación social, Caracas UCAB/UCV, disponible en: <http://www.ucab.edu.ve/eventoscic.html> (Consulta: 2013, mayo 18)
- Prieto, Leticia (2009). **Métodos cuantitativos aplicados a la comunicación social**. Maracaibo: Editorial de la Universidad del Zulia.

Vallejo, B.; Guevara, E. y Medina, S. (2018). Minería de Datos **Revista Científica Mundo de la Investigación y el Conocimiento**. Vol. 2 núm., especial, mayo, ISSN: 2588-073X, 2018, pp. 339-349