

Cartografiar Internetilandia: Una herramienta para el trabajo de campo virtual en Internet

Santiago López-Pavillard

Resumen: Este artículo pretende dar una visión de Internet y de sus usuarios desde una perspectiva holística, mediante su representación visual en forma de mapa. Este mapa puede constituir una buena herramienta de trabajo para el antropólogo que desee realizar análisis cuantitativos y cualitativos de la Red de redes. El artículo muestra un novedoso método para hacer el levantamiento topográfico de Internet desde una perspectiva antropológica, alejado de las representaciones visuales de Internet al uso, basadas en topologías de redes, o en mapas políticos a los que se les ha añadido datos cuantitativos relativos a la Sociedad de la Información (cuantificación de tráfico en bits, de usuarios, de páginas web, de proveedores de acceso a la Red, etc). Se presenta una herramienta que permite tomar decisiones respecto al acceso al campo, pero también para 'crear campo'.

Palabras clave: antropología, etnografía, trabajo de campo, Internet, cartografía, metodología, holismo, información

1. Introducción
 2. Método empleado
 3. Utilidad de esta herramienta de trabajo
 4. Determinación de las formas
 - 4.1 Masas sólidas
 - 4.2 Masas líquidas
 - 4.3 Niveles de interactividad
 5. Determinación de los tamaños
 6. Un primer croquis de Internetilandia
 7. Conclusiones
- Referencias

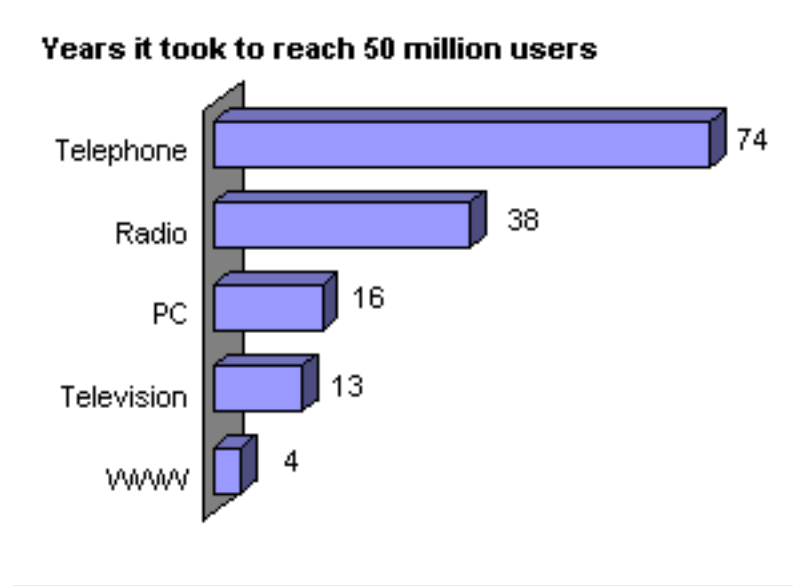
Autor

1. Introducción

En poco más de veinte años trabajando como documentalista en la televisión pública española TVE (<http://www.rtve.es>) he sido testigo de la gran revolución de la gestión de la información en el mundo, gracias a la combinación de informática, telecomunicaciones y digitalización: empezamos usando fichas manuales, luego vinieron los *main-frames* y las primeras bases de datos, llegó la microinformática, se empezaron a mecanizar los registros manuales, luego vino el acceso a Internet, se desarrollaron las Intranets corporativas y, finalmente, se empiezan a digitalizar los documentos primarios. [1]

Hoy en día cualquier usuario, por inexperto que sea, no sólo puede acceder a información referencial contenida en multitud de bases de datos distintas, situadas en cualquier parte del mundo, sino obtener (cada vez más a menudo) una copia del documento original, ya sea textual, gráfico o audiovisual. El papel privilegiado de los *gatekeepers* de la información, sean éstos periodistas o documentalistas, va menguando. Hoy Internet es el mayor medio de comunicación de masas del mundo, con más de 1.000 millones de usuarios¹. Si el teléfono, para

lograr sus primeros 50 millones de usuarios, necesitó 74 años, Internet, siendo un medio de comunicación de mucha mayor complejidad, sólo ha necesitado 4 años para alcanzar esa misma cifra de usuarios². La sociedad asimila con mayor rapidez y naturalidad las nuevas tecnologías que van surgiendo. [2]



Fuente: INTERNATIONAL TELECOMMUNICATION UNION (1999)

El éxito de Internet es rotundo, pero no es un punto final; al revés, es el comienzo de una inmensa revolución acerca de cómo entendemos la información, cuya piedra angular no es el desarrollo de la tecnología en sí sino la capacidad del ser humano por asimilarla cada vez a mayor velocidad. Un aspecto de esta revolución es la concepción del propio sujeto como 'un conjunto determinado de información'. Últimamente la extensión de un ancho de banda ancho (p.e. el ADSL), y la simplificación de uso de muchas aplicaciones de Internet para favorecer la participación en la red (p.e. comunidades virtuales como <http://secondlife.com/> o los blogs <http://www.blogger.com>) están dando lugar a la aparición de una multitud de comunidades virtuales, en las que los individuos se convierten 'en información segmentada y compartida' en función del contexto virtual y el grado de interacción. [3]

Ello nos lleva a preguntarnos qué es para el ser humano 'información', qué hace con ella, por qué cree que la necesita, cómo la manipula, cómo le da vida. En Internet hay muchos mundos, hay muchas vidas, y todas se codifican en bits. Etnografiar al ser humano en Internet nos parece de crucial importancia para entender al ser humano que viene (un *Homo digitalis sapiens*), y que ya está a la vuelta de la esquina. [4]

Este artículo trata de conjugar una larga experiencia en el campo de la gestión de la información con una más recientemente adquirida formación antropológica, para tratar de desarrollar una herramienta de trabajo útil para los científicos sociales. [5]

El objetivo de este texto es presentar un modelo para cartografiar Internet como si se tratara de un espacio físico al que hemos llamado Internetilandia. Es un modelo pensado para la investigación social, que desarrolla tres ideas:

- La importancia de cuantificar en BITS la información que circula por la Red.
- La importancia de caracterizar el tráfico de información en Internet en función de los subprotocolos de comunicaciones usados.
- La importancia de desarrollar un esquema que tipifique los tipos de interacciones entre personas en la Red. [6]

El croquis que presentamos al final de este trabajo, ya lo avanzamos ahora, no nos resulta satisfactorio. Ello se debe a que a pesar de la importancia que tiene Internet en nuestras vidas, y a pesar de ser un medio que puede ser perfectamente escrutado por sistemas de medición y evaluación de todo tipo, carecemos a día de hoy de suficientes datos para desarrollar el modelo que aquí proponemos. [7]

2. Método empleado

En primer lugar, definiremos Internetilandia como cualquier tipo de información que circula por una red informática y que utiliza alguno de los subprotocolos de comunicaciones del *Transfer Control Protocol / Internet Protocol* (TCP/IP). [8]

Para hacer el levantamiento cartográfico de Internetilandia usamos tres tipos de datos:

- Tipo de protocolo usado
- Volumen de información que circula por la red en función del protocolo usado
- Nivel de interactividad que se da entre receptores y emisores de la información [9]

No es fácil determinar el volumen de información que circula por Internet, y menos aún discriminar qué fracción de esta información pertenece a cada uno de los subprotocolos del TCP/IP. Técnicamente la red puede ser perfectamente monitorizada en cualquiera de sus facetas, pero no hemos hallado mejor información cuantitativa de la Red que la que aquí presentamos, que es escasa y poco matizada. [10]

Teniendo en cuenta los tres tipos de datos arriba especificados, hemos seguido los siguientes criterios para determinar las formas y los tamaños contenidos en el mapa (en el punto 4 de este trabajo se desarrollan los criterios propuestos de forma detallada): [11]

CRITERIO DE FORMA:

- División de contenidos y servicios en MASAS SÓLIDAS (información almacenada físicamente en ficheros) y MASAS LÍQUIDAS (información que fluye por la red y que no se llega a almacenar en ficheros de datos estáticos).
- Por MASA SÓLIDA entendemos toda aquella información almacenada en discos duros de ordenadores, cuyos ficheros son susceptibles de ser copiados y almacenados en otros ordenadores.
- Por MASA LÍQUIDA entendemos aquella información almacenada en discos duros de ordenadores o digitalizada a tiempo real y transmitida a la red de Internet, cuyos ficheros no están diseñados informáticamente para ser copiados y almacenados en ordenadores distintos del de procedencia. Este es el caso de los ficheros tipo *streaming* (el

receptor puede reproducir su contenido pero no almacenarlo en su disco duro).

- Subdivisión de la masa sólida en CONTINENTES, y de la masa líquida en OCEANOS y MARES³, siguiendo como criterio de división el tipo de protocolo de comunicaciones usado.
- Subdivisión de los continentes en PAÍSES, y su vecindad (proximidad / alejamiento entre ellos) vendría determinado según los distintos NIVELES DE INTERACTIVIDAD implicados en la generación y comunicación de los distintos tipos de información. [12]

CRITERIO DE TAMAÑO

- El tamaño de las masas sólidas, líquidas y sus subdivisiones viene determinado por el volumen de la información que se mueve dentro de sus fronteras medido en BITS (dado el gran volumen de información existente, la unidad de medida más adecuada, a día de hoy, es el TeraByte (TB)⁴). [13]

3. Utilidad de esta herramienta de trabajo

Como se desprende del título, se pretende establecer una analogía entre Internet y un posible país (Internetlandia⁵), poblado por multitud de habitantes (usuarios), cuya principal actividad es intercambiar y/o difundir información de todo tipo. [14]

Un mapa de Internet levantado bajo estos criterios permitiría:

- Caracterizar a sus habitantes, y las áreas en las que habitan, facilitando al etnógrafo información básica con la que desarrollar una estrategia más adecuada para acceder al campo.
- Propone, además, una innovación al etnógrafo: que, en función de sus intereses, y guiados por el mapa de Internetlandia, escoja la opción de CREAR SU PROPIO CAMPO. Si un etnógrafo quiere obtener la máxima información de su presencia en Internetlandia, seguramente no le quede mejor opción que la de optar a *webmaster*. [15]

Otras ventajas que tiene esta herramienta de trabajo son:

- Permitiría una visión holística de Internet:
 - Combinando datos cualitativos y cuantitativos simultáneamente. Permitiría conocer múltiples características de Internet a simple vista.
 - Diacrónico: permitiría comparar la evolución en el tiempo, pudiendo extraer conclusiones de manera inmediata mediante la comparación de mapas, sólo observando cómo varían superficies, fronteras y costas.
 - Permitiría revelar la complejidad y diversidad de culturas en Internet, dejando de lado las complejidades técnicas de la infraestructura que sustentan la Red.
 - Permitiría contextualizar las subculturas de Internet, determinar homogeneidades y heterogeneidades, por simple vecindad de tipos de información sobre el mapa.

- Permitiría fundamentar razonadamente el procedimiento para triangular la información obtenida en una comunidad virtual acudiendo a otras, cuyo criterio de selección podría objetivarse mediante este modelo de representación. [16]

4. Determinación de las formas

4.1 Masas sólidas

Por MASA SÓLIDA entendemos toda aquella información almacenada en discos duros de ordenadores, cuyos ficheros son susceptibles de ser copiados y almacenados en otros ordenadores. La naturaleza de la información contenida en estos ficheros puede ser cualquiera: vídeo, audio, texto, gráfico, imagen fija, etc. [17]

Los grandes continentes o subprotocolos que forman este tipo de información serían:

- HTTP: Protocolo que gestiona contenidos hipertextuales como son las páginas web. Aquí se incluye la llamada 'blogosfera' (*blogs*).
- POP / IMAP / SMTP: Protocolos que gestionan el correo electrónico. Aquí se incluyen los mensajes generados por listas de distribución tipo Listserv.
- NNTP (*Network News Transfer Protocol*): Protocolo que gestiona los foros de debate, como los USENET.
- P2P (*peer-to-peer*): Protocolo que permite el intercambio de ficheros entre ordenador personales de usuarios.
- FTP (*File Transfer Protocol*): Protocolo que permite el intercambio de ficheros entre servidores y ordenadores personales de usuarios. [18]

4.2 Masas líquidas

Por MASA LÍQUIDAS entendemos aquella información almacenada en discos duros de ordenadores o digitalizada a tiempo real y transmitida a la red de Internet, cuyos ficheros no está diseñados informáticamente para ser copiados y almacenados en ordenadores distintos del de procedencia. Este es el caso de los ficheros tipo *streaming* (el receptor puede reproducir su contenido pero no almacenarlo en su disco duro). [19]

Los grandes océanos y mares que forman este tipo de información son:

- *Voice over Internet Protocol* (VOIP): telefonía por Internet.
- *Internet Group Management Protocol* (IGMP)- *Videostreaming*: IPTV, VoD, WebTV, videoconferencia, etc.
- *Internet Group Management Protocol* (IGMP)- *Audiostreaming*: información procedentes de emisiones de Radio por Internet, o audición de ficheros de audio en formato *streaming*. WebRadio, e-Radio, etc.
- Mensajería instantánea y otras aplicaciones que usan el IRC (*Internet Relay Chat*), así como comunidades virtuales que usan aplicaciones cliente propias bajo el protocolo IP. [20]

El diseño de las fronteras interiores, como el trazado de las costas quedaría en

manos del artista, no, en cambio, el orden de vecindad que se establece. [21]

4.3 Niveles de interactividad

Para nuestros propósitos por INTERACTIVIDAD entendemos la relación comunicativa que se establece entre un emisor y un receptor. El receptor siempre es una persona física, si bien el emisor puede ser una persona física o jurídica. [22]

Consideraremos, por tanto, todos los contenidos emitidos como producidos por una sólo persona (independientemente de si es física o jurídica), si bien el receptor puede ser una persona o varias (un colectivo de personas, es decir, una audiencia). Es decir, la comunicación se establece 1 a 1, ó 1 a muchos; no muchos a 1, o muchos a muchos. [23]

Los criterios para establecer los distintos grados de interacción se basan en dos pares de opuestos:

- Tiempo real / modo diferido
- 1 a 1 / 1 a muchos [24]

Esta primera caracterización de la interacción se matiza con el grado en el que el emisor y el receptor tienen comprometidos sus facultades mentales y sus sentidos en la relación comunicativa:

- Es imprescindible la presencia de ambos para producirse la interacción: p.e., al hablar por teléfono.
- Hace falta la presencia de ambos, pero no de forma simultánea, sino secuencialmente, para que se establezca la relación comunicativa: p.e. en el acto de mandar un mensaje electrónico el proceso comunicativo es secuencial: yo escribo y envío / yo recibo y leo.
- Sólo hace falta un receptor que abra el canal de comunicación: p.e. cualquier contenido de la Masa Sólida de Internet: un emisor publica algo, p.e. un blog, o una canción para ser distribuida mediante la aplicación P2P eMule, pero el receptor no deja de ser más que una hipótesis de la relación comunicativa. [25]

En función de lo antes dicho, ordenamos los distintos subprotocolos en función de un nivel de interacción, considerando que el mayor grado de interactividad lo ofrecería una comunicación 1 a 1 en tiempo real, mientras que en el otro extremo se situaría una comunicación 1 a muchos en modo diferido. [26]

Tabla 1: Asignación de niveles de interacción

Nivel de Interacción	Subprotocolo	Tiempo	Receptor/Emisor
1	VOIP	Simultáneo	1 a 1
2	IRC Comunidades virtuales Mensajería instantánea	Simultáneo	1 a muchos 1 a 1
3	Correo electrónico	Secuencial rápido	1 a 1
4	Foros de debate (Newsgroups)	Secuencial lento	1 a muchos
5	HTTP	Navegación	1 a muchos
6	Video Streaming	Un click - abrir canal	1 a muchos
7	Audio Streaming	Un click - abrir canal	1 a muchos
8	P2P	Un click - abrir canal	1 a muchos
9	FTP	Un click - abrir canal	1 a muchos

5. Determinación de los tamaños

El tamaño de las masas sólidas, líquidas y sus subdivisiones viene determinado por el volumen de la información medido en TERABYTES (TB). [27]

El tamaño de la MASA SÓLIDA depende de la suma del tamaño de todos los ficheros de información almacenado en los discos duros de los servidores conectados a Internet, no del volumen de ancho de banda cuya difusión por la Red generara. [28]

En sentido inverso, el tamaño de la MASA LÍQUIDA depende del consumo de ancho de banda, es decir, del tráfico de bits por la red generado por la difusión de la información comprendida dentro de esta categoría. [29]

Ciertamente, la cuantificación se ha hecho muy *grosso modo*, especialmente en los cálculos del ancho de banda consumido por la Masa Líquida. Los datos aquí presentados se han hallado calculando el ancho de banda generado por un único receptor por canal, estando conectado a ese canal durante toda la emisión. Volvemos a repetir que, sin embargo, lo importante es el modelo de integración e interpretación de datos que ofrece el croquis final, y considerar si puede ser o no una herramienta útil en el trabajo de campo virtual. [30]

Salvo que se señale lo contrario, la fuente usada para cuantificar los distintos tipos de información en Internet es el estudio titulado "How Much Information" (2003), elaborado por los estudiantes del *School of Information Management and Systems*, de la Universidad de California - Berkeley, bajo la dirección de LYMAN Y VARIAN (2003). [31]

Tabla 2: Cuantificación de los distintos tipos de información y asignación de un nivel de interacción [32]

	Subprotocolo		Cuantificación	Nivel interacción
Masa líquida	VOIP		38.221TB ⁶	1
	IRC, mensajería instantánea, comunidades virtuales		274TB	2
	Video Streaming		6.382TB ⁷	6
	Audio Streaming		274TB ⁸	7
Masa sólida	Correo electrónico	C/E	440.606TB	3
		Listservs	563TB	3
	Foros de debate (Newsgroups)		0,27TB	4
	HTTP	Información accesible mediante páginas web recuperables a través de buscadores convencionales tipo Google ('Internet superficial')	167TB	5
		Información NO recuperable a través de buscadores convencionales tipo Google ('Internet profunda')	91.850TB	5
	P2P		5.000TB	8
	FTP		97TB	9

6. Un primer croquis de Internetilandia

Tabla 3: Determinación del número de unidades (casillas), en función de la cuantificación resultante (1 casilla = 5.834,34 TB) [33]

TIPO DE INFORMACIÓN		TB	% del total de TB	Nº casillas
VOIP		38.221	6,55	29
IRC, mensajería instantánea, comunidades virtuales		274	0,05	0
Video Streaming		6.382	1,09	5
Audio Streaming		274	0,05	0
Correo electrónico	C/E	440.606	75,52	340
	Listservs	563	0,10	0
Foros de debate (Newsgroups)		0,27	0,00	0
HTTP	Información accesible mediante páginas web recuperables a través de buscadores convencionales tipo Google ('Internet superficial')	167	0,03	0
	Información NO recuperable a través de buscadores convencionales tipo Google ('Internet profunda')	91.850	15,74	71
P2P		5.000	0,86	4
FTP		97	0,02	0
Total TB		583.434	100,00	450,00
Total casillas		450		

Leyenda: [34]






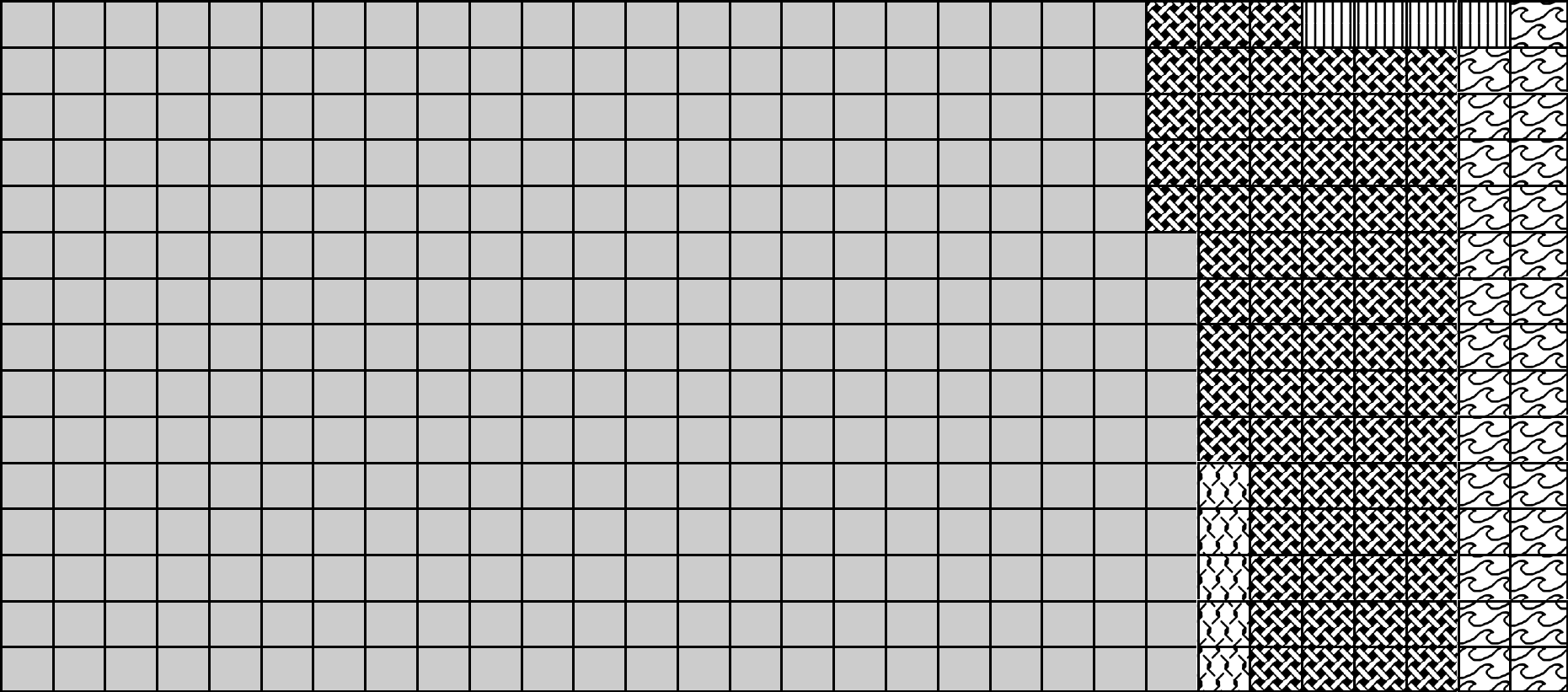
	Correo electrónico
	Internet profunda o invisible
	VOIP
	P2P
	Videostreaming

Tabla 4: Croquis de Internetilandia (450 casillas)
Escala 1 casilla: 5.834,34TB [35]



7. Conclusiones

- A pesar de la gran importancia de Internet, se carece de suficientes datos acerca del tráfico de información por la red, y su caracterización técnica (p.e. cuánto ancho de banda corresponde a cada subprotocolo de la gran familia de protocolos TCP/IP. Muestra de esta carencia son los proyectos que están desarrollando en este sentido la *National Science Foundation* norteamericana junto con actores del sector privado a través de CAIDA⁹ (*Cooperative Association for Internet Data Analysis*) [36]
- Mientras no exista una buena caracterización del tráfico de información por la red Internet no se podrá desarrollar una adecuada caracterización de Internet como Internetilandia, tal como se propone en este artículo. [37]
- A pesar de los datos escasos, y a veces, difíciles de interpretar, sí podemos observar en el croquis aquí presentado que Internet presenta, desde el punto de vista cuatitativo, grandes desequilibrios que es necesario compensar desde el punto de vista gráfico, yendo tal vez a escalas logarítmicas. [38]
- Usando una escala de 1 unidad gráfica (p.e. una casilla) = 5.834,34 TB, hace que mucha información de valor cualitativo en Internet no se vea reflejada en el mapa, como se puede apreciar en la tabla del punto n°6, donde algunos apartados tiene asignados 0 casillas. [39]
- Creemos que para el trabajo etnográfico en Internetilandia no se necesita contar con un mapa de pequeña escala. El etnógrafo, en función de la especialización de su trabajo de campo, siempre podrá levantar planos detallados de su área de interés siguiendo este método simple que aquí se propone. [40]

Referencias

International telecommunication Union (1999): "Challenges to the Network: Internet for Development, 1999", en <http://www.itu.int/ITU-D/ict/publications/inet/1999/ExeSum.html> [February 12, 2007]

Lyman, Peter and Hal R. Varian (2003): "How Much Information", en <http://www.sims.berkeley.edu/how-much-info-2003> [6 de febrero de 2007].

Autor

Santiago LÓPEZ PAVILLARD es antropólogo, y trabaja como documentalista en TVE. Sus preferencias de investigación son el estudio del concepto de 'Información' desde la perspectiva antropológica, y el estudio de los diferentes Estados de Consciencia y su relación con la cultura humana. Posee el Diploma de Estudios Avanzados por el Departamento de Antropología Social de la Facultad de Ciencias Políticas y Sociología, de la Universidad Complutense de Madrid.

Contacto:

Santiago López-Pavillard

Correo electrónico:

slpavillard@cps.ucm.es

Sitio web:

<http://emc.mercurialis.com>

Fecha de composición del texto: 14 de febrero de 2007

Revisión de datos personales: 9 de noviembre de 2011

NOTAS

1) Fuente: <http://www.internetworldstats.com/stats.htm> [February 14, 2007]

2) Gráfico - "Challenges to the Network: Internet for Development, 1999", October 1999, Updated for Telecom 99 - International telecommunication Union. Figura 1: Internet growth: No longer exponential, but still mightily impressive. Fuente del gráfico: ITU World Telecommunication Indicators Database, Internet Software Consortium, Compaq, RIPE., en <http://www.itu.int/ITU-D/ict/publications/inet/1999/ExeSum.html> [February 12, 2007]

3) Para tecnologías híbridas no tenidas en cuenta en este trabajo, podríamos pensar en añadir al mapa accidentes geográficos como ríos y montañas. Por ejemplo, el sitio YouTube.com podría ser representado como un gran río Amazonas discurriendo entre la gran masa sólida de información HTTP. En este trabajo, sin embargo, consideramos todo el flujo de información generado por el *videostreaming* de YouTube y sitios similares dentro de la MASA LÍQUIDA.

4) Un TB son 1.024 GB. Es importante distinguir entre bits (b) y bytes (B). 1B equivale a 8b.

5) La palabra *Internetilandia*, a pesar de ser un término que puede surgir fácil y espontáneamente al pensar acerca de etnografía e Internet, da 0 resultados si la buscamos en <http://www.google.com>. La palabra *Internetland* aparece en 56.200 documentos [February 14, 2007]

6) Hemos tenido en cuenta el consumo anual telefónico (114.000 millones de minutos), y teniendo en cuenta las previsiones que señalaban que para el año 2007 el 75% de dicho consumo se haría por IP, hemos multiplicado el consumo resultante por 64Kbps.

7) Tomando como referencia el buscador <http://wwitv.com>, [February 12, 2007] que contiene 2.318 referencias de emisoras de televisión por Internet, hemos considerado una emisión media diaria anual de 12 horas, con un bitrate de 1.500 Kbps.

8) Hemos considerado 585 emisoras de radio por Internet, emitiendo las 24h. Del día, 365 días al año, con un bitrate de 128 kbps.

9) URL de CAIDA: <http://www.caida.org>

Los objetivos de CAIDA son:

- * Encourage the creation of Internet traffic metrics (in collaboration with IETF/IPPM and other organizations); and work with industry, consumer, regulatory, and other representatives to assure their utility and universal acceptance.

- * Create a collaborative research and analytic environment in which various forms of traffic data can be acquired, analyzed, and (as appropriate) shared.

- * Foster the development of advanced methodologies and techniques for: traffic performance and flow characterization, simulation, analysis, and visualization. Specific areas of future impact include real-time routing instability diagnosis and evolution for next generation measurement and routing protocols (multicast and unicast).

(En <http://www.caida.org/home/about/> [February 14, 2007])