

La información en Internet: Breve estado del arte para discutir el poder los usuarios v/s los medios tradicionales de comunicación en la red

Information on the Internet: A quick view to state of art to discuss the users's power v/s traditional communication medias on the network

DIEGO SÁEZ-TRUMPER*

Resumen

El siguiente artículo es un breve resumen de investigaciones recientes sobre los fenómenos de propagación de información a través de Internet y sobre el rol que cumplen los usuarios frente a las versiones "on-line" de los grandes medios de comunicación. Con este propósito, damos un vistazo general a las técnicas utilizadas en las Ciencias de la Computación

* Departamento de Tecnologías de la Información y las Comunicaciones, Universidad Pompeu Fabra, Barcelona, España. Edifici Tànger, Campus de la Comunicació-Poblenou, Tànger 122-14008018. E-mail: diego.saez@upf.edu.

para analizar dichos procesos. En definitiva, el artículo permite familiarizar al lector con las investigaciones realizadas desde este campo sobre los procesos de difusión de información en redes, introduciendo conceptos y técnicas claves para nuestra disciplina.

Palabras clave: redes sociales, medios sociales, difusión de información, computación.

Abstract

This article is a short summary of state of the art about information propagation over the Internet and the role of the users and the mainstream media over that process. This article tries to introduce the techniques and findings to analyses the information diffusion process from the point of view of Computer Scientists. Summarizing this article allows the reader to get familiar with big picture of the computer scientists about information diffusion.

Key words: social networking, social media, dissemination of information, computing.

Introducción

Se estima que en 2011 había más de 150 millones de blogs y 200 millones de cuentas de Twitter. Por un lado, esto significa una enorme diversidad en la oferta informativa. Así los *medios sociales en Internet*¹ son una oportunidad para

¹ Medios sociales (del inglés *social media*) es un concepto que abarca las redes sociales, blogs, microblogs y otras plataformas para el intercambio de información a través de Internet. "Los medios sociales son entornos que propician el incremento masivo de la publicación de información y opinión en la red por parte de los usuarios, así como su creciente participación en redes horizontales de comunicación". En línea, disponible en: <http://www.cenitsocialmedia.es/>.

individuos o pequeños colectivos que desean difundir informaciones a un costo mínimo. Por otro lado, los medios tradicionales (diarios, canales de televisión, etc.) también ocupan la red para difundir sus contenidos. Se produce, por lo tanto, una competencia por la “cyber audiencia” entre los usuarios (individuos o colectivos) y los medios de comunicación tradicionales. ¿Quién gana esta competencia?, ¿puede realmente un bloggero competir con CNN?, o mejor dicho, ¿pueden muchos bloggeros, twitteros, etc., competir con CNN?, ¿es realmente la Internet un espacio democratizador o es tan sólo una ilusión de horizontalidad?

Para responder estas preguntas podemos remitirnos a estudiar la enorme cantidad de datos que cada acción –tanto de los usuarios como de los medios tradicionales– van dejando en la Internet. Si bien ya desde los años 60 disciplinas como la Economía y la Sociología han estudiado los procesos de difusión de información, en la actualidad, la gran cantidad de información disponible, así como las dificultades técnicas para procesarla, han hecho que temas como estos cobren gran popularidad entre físicos, matemáticos, ingenieros y especialistas relacionados con la computación, quienes, utilizando técnicas de Inteligencia Artificial y Minería de Datos, pueden dar algunas respuestas –y nuevas preguntas– a estas interrogantes. Es por ello que este artículo intenta entregar una mirada panorámica de las técnicas utilizadas desde esta área del conocimiento.

Los usuarios versus los grandes medios

La Internet se presenta como una gran oportunidad para democratizar el acceso

a la información, así como también una oportunidad de compartir y publicar contenidos. Cualquier persona con mínimos conocimientos informáticos puede crear un medio de difusión para sus ideas/opiniones/preocupaciones. Sin embargo, es conocido que la popularidad tanto de las páginas, como de los usuarios de los medios sociales responden a una *ley de potencias* (Mislove *et al.* 2007), es decir, que unas pocas fuentes de información concentran el mayor número de seguidores/lectores. Más aún, algunos autores plantean que la concentración en Internet es mayor que en que el mundo “off-line”: “Internet provee a cualquier ciudadano de una audiencia potencial de miles de millones, de la misma forma en que potencialmente cualquier persona puede ganar la lotería. En su entusiasmo, muchos olvidan sacar las cuentas, y los números muestran que las posibilidades de tener éxito a lo grande, en la red, son extremadamente pequeñas” (Hindman Cit. en Castillo 2010, en línea).

Basado en estadísticas que comparan la cantidad de diarios vendidos por periódicos grandes y pequeños versus el número de visitas que reciben sus páginas en Internet, Hindman concluye que la concentración es mayor en el último caso. Este comportamiento puede ser explicado por el fenómeno “conexión preferencial” (*preferential attachment*) -también conocido como “el rico se hace más rico”-, según el cual en una red aquellos nodos (usuarios) que tienen mayor número de aristas (seguidores) presentan, a su vez, una mayor probabilidad de crecer respecto a aquellos con un bajo número de seguidores. Varios estudios demuestran que este fenómeno se produce en general en Internet (Barabási y Albert 1999), por lo cual es fácil entender que los grandes medios de comunicación pueden tener como punto de

partida un número mayor de seguidores que un usuario, por lo cual su número de seguidores crecerá en el futuro por dentro de la red. En otras palabras, su popularidad en el mundo “off-line” les ayuda a ser populares en la red.

Sin embargo, un gran número de seguidores no significa necesariamente ser más influyente o lograr esparcir más información. En el artículo “La falacia del millón de seguidores” (Cha *et al.* 2010), se muestra que las cuentas con más seguidores en Twitter no suelen ser las más influyentes. Por ejemplo, celebridades como “Lady Gaga” tienen un gran número de seguidores en Twitter pero el impacto de lo que publican es bajo, mientras otros usuarios anónimos, con menor número de seguidores, son capaces de generar cascadas de información que alcanzan a un gran número de usuarios (Saez-Trumper *et al.* 2012).

Esto último se debe a que la popularidad de una noticia o contenido de Internet es bastante más compleja de determinar que el impacto de una fuente individual. Esto es así ya que una característica inherente a la comunicación en red está dada por los procesos de “contagio”, es decir, cómo un nodo pasa información a otro, y éste, a su vez, lo retransmite a sus vecinos, generando así procesos de influencia indirecta. Existen diversos modelos y aproximaciones para estudiar este fenómeno. Una de ellas consiste en agrupar las fuentes según su naturaleza, por ejemplo, en lugar de estudiar a un bloggero o a un canal de televisión determinado, se pueden agrupar todos los bloggeros, todos los canales de televisión, y ver cómo se relacionan entre ellos. Esto fue lo que hicieron Yang y Leskovec (2011) en un estudio donde examinaron más de 170 millones de noticias que circulaban por la Internet, obtenidas durante 8 meses desde

sitios web de periódicos, canales de televisión, agencias de noticias y diversos tipos de blogs. A partir de esto agruparon cada fuente de información según su naturaleza y analizaron las curvas de popularidad de las noticias, mediante un estudio de la relación temporal entre distintos tipos de fuentes. Más específicamente, su investigación consistió en ver cómo ciertos pequeños fragmentos de texto, particularmente citas textuales llamadas *memes*², se repetían en diferentes medios. Luego, adaptando un algoritmo de agrupamiento (*clustering*), obtuvieron las distintas formas de la curva de popularidad de las noticias, y qué tipo de medio explicaba esa popularidad.

Como se puede ver en la figura 1 existen noticias donde bloggeros se adelantaron a los medios tradicionales (*clusters* 5 y 6). Si bien estos *clusters* agrupan sólo un 16% de las noticias estudiadas, no deja de ser interesante que ciudadanos comunes logren adelantarse y/o marcar la pauta de los grandes medios de comunicación. Sobre todo, resulta relevante de considerar dado que estos datos fueron obtenidos antes del 2010, por lo que es probable que esta tendencia haya aumentado junto con el constante crecimiento de los medios sociales. Las formas de las curvas están también relacionadas con el rol que cumplen las diferentes partes. Así las que tienen origen exógeno, es decir, su crecimiento se debe a causas externas a la red, presentan un rápido incremento y caída, mientras que las de origen endógeno tienen una subida y caída más suave. Esto en la práctica significa que a los *bloggeros*

² La palabra “meme” hace referencia al concepto introducido por Richard Dawkins en su libro *El gen egoísta* (2000), definida como la mínima unidad de información cultural que puede ser transmitida de una persona a otra haciendo un paralelismo con los genes.

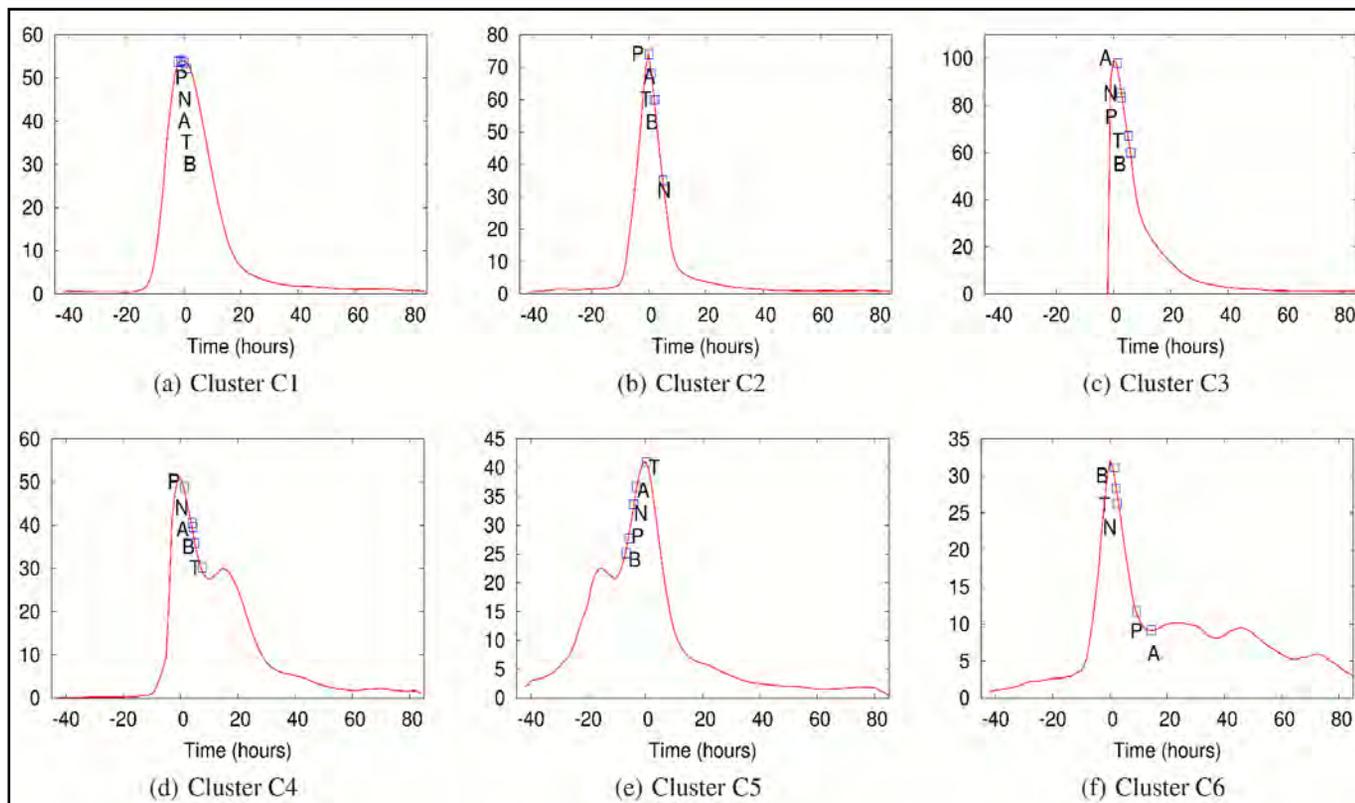


Figura 1. Curvas de popularidad de las noticias en Internet y sus fuentes. B: bloggers, P: blog profesional, N: periódico, A: agencias de noticias, T: canales de televisión (Yang y Leskovec 2011).

les toma más tiempo lograr popularizar una noticia, en comparación con los grandes medios de comunicación.

Conectores, genios y vendedores

Si en determinados casos los usuarios normales logran hacer popular un contenido, es interesante preguntarse cuáles son los mecanismos detrás de este fenómeno. En la actualidad existen diversos estudios intentando predecir qué contenido será popular en el futuro y cómo encontrar los usuarios más influyentes en determinados temas. Las aplicaciones de estas investigaciones en campañas publicitarias y políticas es evidente. Sin

embargo, hasta hoy, no existe un modelo único que pueda describir los procesos de traspaso de información. Diferentes conceptos son utilizados para intentar modelar estos procesos. Una idea importante es la de innovadores versus imitadores. Ésta nos sugiere que sólo unos pocos sujetos están dispuestos a innovar y ser los primeros en proponer algo (o hablar sobre algo nuevo), no obstante, existe una mayoría que espera a que exista un cuerpo de opiniones validadas para “contagiarse” o imitar iniciativas (o simplemente hablar de un tema ya instalado). Esta propuesta, tomada de modelos económicos sobre adopción de productos (Bass 1969), señala que existe una relación (descrita por una ecuación diferencial) entre la cantidad de innovadores e imitadores.

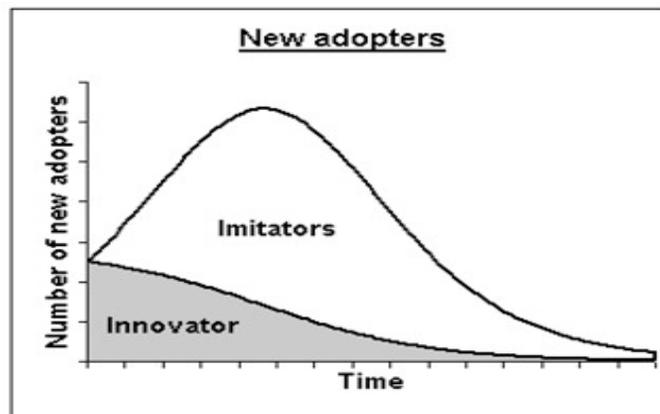


Figura 2. El modelo de Bass (1969) sobre innovadores e imitadores. Muestra la relación temporal existente entre ambos.

Luego, la pregunta natural es cómo alcanzar la cantidad necesaria de innovadores (masa crítica) necesaria para desatar una reacción en cadena. Según Malcom Gladwell (2000), esto se consigue cuando se da la adecuada combinación entre una serie de tipos personalidad dentro de una red. Para Gladwell existen al menos tres tipos importantes de usuarios que impulsan la popularidad de un contenido: los genios (*mavens*), los conectores y los vendedores. En primer lugar, los genios son aquellos que poseen alguna información privilegiada o de primera fuente; en segundo lugar, los conectores son quienes están centralmente ubicados en las redes sociales, teniendo un gran número de “amigos” o “seguidores”; y, en tercer lugar, los vendedores o “evangelistas” son aquellos que se dedican sistemáticamente a difundir un contenido.

Otros autores consideran que lo más relevante es el contenido de la información que se está transmitiendo. Así, en su libro *Grouped* (Agrupados) Paul Adams (2011), establece que los contenidos que son capaces de expandirse ampliamente por las redes sociales, son aquellos con características sorprendentes y generadores de conversaciones o debates al interior de los pequeños grupos de amigos que cada usuario tiene. Para Adams no hay usuarios influyentes, sino una serie de pequeñas comunidades que son conectadas entre sí por los usuarios. En esa perspectiva, cada usuario es influyente dentro un determinado marco, ya que cada uno de ellos posee una red social única. Luego, cada usuario puede “contagiarse” dentro de un círculo, y pasar esa información hacia otro.

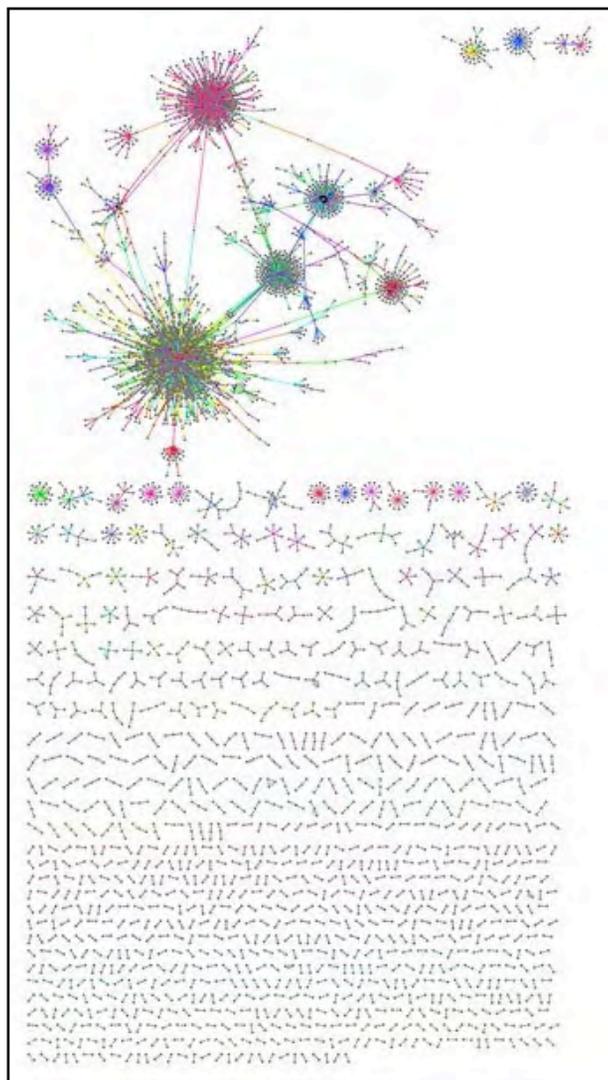


Figura 3. Representación gráfica del contagio de una noticia a través de Twitter (Kwak *et al.* 2010). En la figura se observan distintos patrones de propagación de una misma noticia. Por un lado, existe un gran componente conectado (parte superior de la figura) y, por otro, una serie de componentes pequeños (islas hacia el final de la figura).

Vemos los que queremos ver

La enorme cantidad de redes sociales y blogs debiesen permitir una mayor diversidad en las informaciones a las que se puede acceder, en contraposición al acceso que permiten los medios tradicionales. Sin embargo, esto no necesariamente significa que los lectores busquen una mayor diversidad de opiniones sino muy por el contrario. Lo que sucede, en la práctica, es que las personas suelen buscar informaciones que confirmen sus opiniones y creencias, por lo cual tienden a encerrarse en grupos pequeños, fuertemente conectados entre sí y con escasa asociación con otros de

posiciones diferentes. Un estudio sobre los blogs políticos durante la campaña presidencial de 2004 en Estados Unidos, mostró que los blogs de liberales y conservadores tendían a conectarse y citarse entre ellos, ignorando a los del rival. Estudios similares en Twitter (An *et al.* 2011) han confirmado esta tendencia.

Si bien no es novedoso que la gente tienda a buscar información y a citar aquellos que piensan de una manera similar, lo que sí es novedoso es que se han podido estudiar los cambios de opinión y los movimientos de las comunidades en tiempo real, utilizando herramientas como el análisis de sentimientos³.

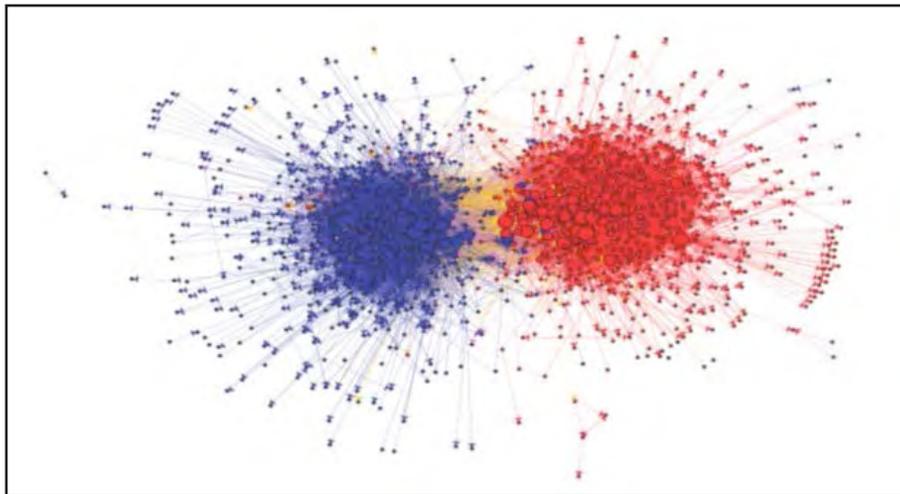


Figura 4. Estructura de la comunidad de *blogs* de políticos en las elecciones presidenciales de Estados Unidos (2004). El color del nodo representa la orientación política: el azul corresponde a los conservadores y el rojo a los liberales. Las aristas naranjas representan una conexión desde los liberales hacia los conservadores, y las aristas violetas representan la conexión inversa. El tamaño de los nodos da cuenta de la cantidad de enlaces apuntando hacia ellos (Adamic y Glance 2005).

³ Análisis de Sentimientos (*Sentiment Analysis*) también conocido como “Minería de Opinión” (*Opinion Mining*) es una técnica que, a través del Procesamiento del Lenguaje Natural (NPL del inglés *Natural Language Processing*) y el Aprendizaje de Máquina (*Machine Learning*), busca extraer de manera automática el contenido

subjetivo de un texto. Con esto se puede, por ejemplo, analizar en poco tiempo las opiniones publicadas en Internet sobre un determinado acontecimiento. Existe una extensa línea investigativa aplicando técnicas como el análisis de sentimientos en diversas áreas como política, revisión de productos o detección de epidemias.

Este hecho ha llevado, entre otras cosas, a que algunos científicos afirmen que las redes sociales se pueden utilizar para predecir el resultado de elecciones presidenciales, alcanzando márgenes de error similares a las encuestas tradicionales (Watts 2011). Otros niegan que esto sea posible, argumentando que existe un claro sesgo ya que los usuarios de redes sociales corresponden a un subgrupo de la población muy delimitado. No obstante, debido a su bajo costo y al creciente número de usuarios, es muy probable que este tipo de herramientas de predicción se perfeccionen y cobren popularidad en el futuro.

Consideraciones finales

Hemos visto que en la mayoría de las informaciones que circulan por Internet, los medios tradicionales siguen teniendo un mayor poder que los usuarios, sin embargo, también hemos constatado que a través de los procesos de contagio la información puede difundirse al interior de las redes sociales y, en determinadas circunstancias, los usuarios comunes pueden adelantarse y marcar pauta a los grandes medios.

Lo anterior, sumado al hecho de que los usuarios tienden a buscar y enlazar información que confirme sus posiciones, lleva a suponer que la Internet está amplificando una serie de fenómenos que existen en el mundo real. Pensar en las redes sociales como un espacio

idílico exento de problemas exteriores, es engañarse. Las redes y medios sociales son reflejo de una parte de la sociedad, y son también herramientas que facilitan la comunicación: nada más que eso, pero tampoco nada menos. Las oportunidades para difundir información, así como estudiar estos procesos es algo que resultaba inimaginable hace décadas atrás. El *poder de los usuarios* constituye, por tanto, una revelación inédita.

Para finalizar, cabe mencionar ciertas consideraciones sobre la privacidad y otros usos (o abusos) vinculados a la web. Cuando se habla de privacidad en Internet es común tomarlo desde el punto de vista individual, es decir, cómo se afecta la intimidad del individuo y cómo ésta puede ser rastreada en la red. Existen preocupaciones fundadas sobre el seguimiento que, por ejemplo, un gobierno determinado pueda realizar sobre alguno de sus detractores⁴. Sin embargo, esto no es novedoso; el espionaje y la inteligencia militar/policial son algo muy antiguo, y sin duda su versión en digital es tan sólo un capítulo más en esa historia. Lo realmente novedoso no es el seguimiento a nivel personal, sino *el análisis del comportamiento colectivo* que se puede realizar con las nuevas tecnologías. En este punto es donde están las nuevas oportunidades para la investigación científica, así como la amenaza de que estas informaciones queden sólo en unas pocas manos. La herramienta está ahí y el efecto que ésta pueda tener estará dada por quien la aprenda a utilizar mejor. Los usuarios tienen su oportunidad.

⁴ En el "caso "Wikileaks" tribunales estadounidenses solicitaron información privada de determinadas cuentas de Twitter.

Bibliografía

- Adamic, L.; Glance, N. 2005. "The political blogosphere and the 2004 U.S. election: divided they blog". *Proceedings of the 3rd international workshop on Link discovery (LinkKDD '05)*. ACM. New York, EE.UU.
- Adams, P. 2011. *Grouped: How small groups of friends are the key to influence on the social web*. Berkeley: New Riders.
- An J. *et al.* 2011. "Landscape in Twitter: A world of new conventions and political diversity". *Proceedings of the 5th ICWSM*. AAAI. Barcelona, España.
- Barabási, A.L.; Albert, R. 1999. "Emergence of scaling in random networks". *Science* 286: 509–512.
- Bass, F. 1969. "A new product growth model for consumer durables". *Management Science* 15, 5: 215–227.
- Castillo, C. 2010. "El Mito de la Democracia Digital". *Manzana Mecánica*. En línea, disponible en: http://manzanamecanica.org/2010/04/el_mito_de_la_democracia_digital.html (visitado el 20 de marzo de 2012).
- Cha, M. *et al.* 2010. "Measuring User Influence in Twitter: The Million Follower Fallacy". *Proceedings of the Int'l AAAI Conference on Weblogs and Social Media (ICWSM)*. AAAI. Washington DC, EE.UU.
- Dawkins. R. 2000. *El gen egoísta*. Barcelona: Salvat.
- Gladwell, M. 2000. *The Tipping Point: How Little Things Can Make a Big Difference*. New York: Little Brown.
- Kwak, H. *et al.* 2010. "What is Twitter, a Social Network or News Media?". *19th World-Wide Web (WWW) Conference*. ACM. Raleigh, EE.UU.
- Mislove, A. *et al.* 2007. "Measurement and Analysis of Online Social Networks". *Proceedings of the 5th ACM/USENIX (IMC'07)*. ACM. San Diego, EE.UU.
- Sáez-Trumper, D. *et al.* 2012. "Finding Trendsetters in Information Networks". *Proceedings of KDD Conference*. Beijing, China.
- Watts, D. 2011. "The Dream of Prediction: Why You Should Be Skeptical". *Yahoo! News*. En línea, disponible en: <http://yearinreview.yahoo.com/2011/blog/8569/predictions-why-you-should-be-skeptical>. (vistado 20 de marzo de 2012).
- Yang, J.; Leskovec, J. 2011. "Patterns of Temporal Variation in Online Media". *ACM International Conference on Web Search and Data Mining (WSDM)*. Hong Kong, China.

