

Litterio, Mario; Nantes, Esteban Alberto

REDES SOCIALES ONLINE: ANÁLISIS DE TEXTOS PARA GENERAR CONTENIDOS PARA DISTINTOS SEGMENTOS

XXXV Encuentro de Docentes Universitarios de
Comercialización de Argentina y América Latina

26, 27, 28 y 29 de octubre de 2021

Litterio, M., Nantes, E. A. (2021). Redes sociales online: análisis de
textos para generar contenidos para distintos segmentos. XXXV
*Encuentro de Docentes Universitarios de Comercialización de
Argentina y América Latina. En RIDCA. Disponible en:*
<https://repositoriodigital.uns.edu.ar/xmlui/handle/123456789/6046>



Esta obra está bajo una Licencia Creative Commons
Reconocimiento-NoComercial-Sin Derivados 4.0 Internacional (CC BY-NC-ND 4.0)
<https://creativecommons.org/licenses/by-nc-nd/4.0/>



XXXV ENCUENTRO DE DOCENTES
UNIVERSITARIOS DE COMERCIALIZACIÓN DE
ARGENTINA Y AMÉRICA LATINA

Universidad Católica
de Santiago del Estero
Scientia Deo Et Patriae Servire



REDES SOCIALES ONLINE: ANÁLISIS DE TEXTOS PARA GENERAR CONTENIDOS PARA DISTINTOS SEGMENTOS

XXXV EDUCA-AL

**Área temática:
Trabajos de investigación.**

**Autores:
Litterio, Arnaldo Mario.
Nantes, Esteban Alberto.**

**Cátedras:
Comercialización - Comercialización LA - Seminario de Investigación de Mercados.
Departamento de Ciencias de la Administración.
Universidad Nacional del Sur.**

**E-Mails:
litterio@uns.edu.ar
enantes@uns.edu.ar**

RESUMEN

El presente trabajo da continuidad a la línea de investigación sobre la influencia en redes sociales online que se inició a partir de la publicación *Marketing and social networks: a criterion for detecting opinion leaders*, realizada por este mismo equipo en el *European Journal of Management and Business Economics* y propone un método para mostrar cómo a través de herramientas como NodeXL se pueden interpretar mapas de audiencia e interacciones en una red social como Twitter partiendo de la elección de un hashtag que represente el mercado de interés.

En concreto se muestra a través del método del caso la forma de analizar grupos o clusters consolidados de individuos de una red, determinar su influencia y caracterizarlos a través de diferentes métricas, incluyendo los léxicos que utilizan. Esto ayuda a ponderar y elegir aquellos subgrupos de interés para maximizar el potencial del mensaje, generar la conversión deseada, ya sea lograr una venta, posicionar una idea o conseguir tráfico hacia un sitio web o red social, entre otros.

Para la extracción de la red social, análisis y visualización se utilizan los paquetes de software NodeXL y Excel.

La citada metodología extiende la propuesta de trabajos anteriores ya que muestra un método de trabajo para lograr mayor eficacia y eficiencia en el marketing de redes sociales. El análisis del texto y hashtags relevantes proporcionan una información cualitativa valiosa para la generación de acciones de marketing y de contenidos, pudiendo mejorar el alcance y efectividad de las acciones en redes sociales online.

PALABRAS CLAVE

Análisis de redes sociales, marketing digital, influenciadores, mapa de audiencia, minería de texto

KEYWORDS

Social network analysis, digital marketing, influencers, audience map, text mining.

INDICE

1. INTRODUCCIÓN
2. MARCO TEÓRICO
3. OBJETIVO
4. METODOLOGÍA
5. CONCLUSIONES

1- INTRODUCCIÓN

El análisis del consumidor, el entorno competitivo, la gestión del conocimiento, el posicionamiento y la promoción de productos y servicios en el mundo online son funciones claves del marketing moderno.

En la actualidad, el mapa sobre el que organizaciones comerciales, políticas y no gubernamentales desarrollan su actividad debe ser trazado sobre un terreno en el que se entrelazan las disciplinas del marketing y el análisis de las redes sociales. Esta tendencia se consolidó y aumentó por el impacto de la pandemia de Covid-19 que encerró a las personas y las llevó a socializar, educarse, informarse y consumir a través de estas plataformas tecnológicas.

Las redes sociales online facilitan el diálogo social en línea entre personas, marcas e instituciones, proveyendo una plataforma en la que opiniones y comportamientos de los consumidores se ven afectados por complejas redes de influencia social.

Las redes sociales online se convierten en un nuevo terreno en cual las marcas y empresas deben redefinir su relación con sus consumidores (Benedetti, 2015).

El presente trabajo da continuidad a la línea de investigación sobre la influencia en redes sociales online que se inició a partir de la publicación *Marketing and social networks: a criterion for detecting opinion leaders*, realizada por este mismo equipo en el *European Journal of Management and Business Economics*.

En trabajos anteriores, entre los cuales se encuentra la publicación ya mencionada, se hizo un primer abordaje al tema desde las herramientas que provee el marketing y la teoría de redes sociales, concretamente, el uso de métricas de centralidad. Se logró demostrar que es posible el uso del análisis de redes sociales aplicado a comunidades online para detectar influyentes en el marco del marketing digital. El marketing tradicional se complementa con la posibilidad de operar directamente sobre estos individuos y generar un efecto multiplicador basado en el boca a boca digital.

El resultado fue el desarrollo de una herramienta para analizar las redes sociales online que permite caracterizar y detectar a los actores influyentes a partir de su posición relativa dentro de una red social, determinada por una combinación de su centralidad de autovector y centralidad de intermediación.

En sucesivas publicaciones, que incluyeron métodos como la simulación basada en agentes, la minería de texto, el análisis sobre redes bimodales y la experimentación sobre diferentes comunidades nucleadas en redes sociales no abordadas con anterioridad, se validó la utilidad de la herramienta para encontrar actores en esas redes que son difusores eficientes de información, pudiéndose detectar de manera más rápida y eficiente en comparación con otros métodos.

En la misma línea de investigación, se aplicó el método propuesto a un caso real en el sector privado. El resultado del experimento fue exitoso, tanto por los resultados obtenidos como por el hecho de que la herramienta fue bien recibida y entendida por el cliente,

entendiendo el objetivo, funcionamiento y utilidad para incorporarla a las acciones de marketing.

Todas estas investigaciones sirvieron para validar y robustecer el modelo propuesto originalmente, demostrando su funcionamiento, su aplicabilidad en redes diferentes sociales a nivel teórico y en el campo a través de la aplicación a un caso testigo del mundo real.

Continuando con la línea de investigación, en el entendimiento de que la propuesta no deber terminar en la detección del influyente y que a su vez este es un actor clave de una comunidad sobre la que se quiere lograr un resultado, se propone mostrar cómo a través de herramientas como NodeXL se pueden interpretar mapas de audiencia e interacciones en una red social como Twitter partiendo de la elección de un hashtag que represente la comunidad de interés.

Este trabajo está organizado de modo que en la primera parte se elabora un marco conceptual que incluye títulos de marketing y el análisis de las redes sociales, incluyendo los procesos de influencia en la decisión de compra, el estudio y análisis de redes sociales en general y redes sociales online en particular. Se introducen además los métodos y técnicas utilizadas para obtener y procesar la información obtenida de Internet.

En la segunda parte, se detalla el objetivo del trabajo.

Seguido, se procede a introducir el caso testigo para el estudio y desarrollar la metodología propuesta para el análisis de redes sociales.

Finalmente, se ofrecen las conclusiones a las que se arribó y sus limitaciones, en busca de promover la discusión y posibles extensiones a este trabajo.

2- MARCO TEÓRICO

El boca a boca y el marketing online

Tradicionalmente la comunicación de las marcas con sus potenciales consumidores se ha realizado en forma unidireccional, dejando fuera de cuestión la interacción con los consumidores, rol que ha quedado reservado, en el mejor de los casos, a la investigación e inteligencia de mercados (Bacile, Ye & Swilley, 2014; Benedetti, 2015).

El boca a boca o *word of mouth*, ha sido extensivamente estudiado y reconocido como un factor de influencia clave en las decisiones de compra de los consumidores (Lang & Hyde, 2013; King, Racherla & Bush, 2014), y tiene un efecto positivo en la relación del consumidor con la marca y en otros resultados del marketing (Hudson, Huang, Roth & Madden, 2016; Wang & Gon Kim, 2017).

Hewett y otros (Hewett, Rand, Rust & van Heerde, 2016) postulan que la naturaleza de la comunicación de marca ha cambiado con el advenimiento de las tecnologías online, y cuantifica su efecto en el sentimiento del consumidor y resultados de negocio.

Hoy el proceso de compra de un producto o servicio está fuertemente permeado por la experiencia del potencial consumidor dentro del ecosistema de Internet en general y las redes sociales en particular. A través de la web se pueden descubrir productos, analizar y comprar alternativas, ver videos de su uso, leer opiniones y concretar la transacción.

La comunidad de potenciales consumidores de una marca comienza a dialogar entre sí y con la marca en lugar de ser simple receptora de un mensaje, y esta interacción afecta profundamente sus percepciones y decisiones de compra (Benedetti, 2015).

Las redes sociales online tienen un denominador común que es que dependen del contenido generado por los propios usuarios, el cual generalmente se encuentra relacionado con marcas y tiene el potencial de influenciar la percepción de la marca por parte de los consumidores (Smith, Fischer & Yongjian, 2012; Nam & Kannan, 2014).

El "apego" a los medios sociales online está relacionado positivamente con conductas de consumo y apoyo a las marcas, lo que convierte a algunas personas en objetivos naturales de iniciativas de marketing (VanMeter, Grisaffe & Chonko, 2015).

Varios estudios previos señalan una mayor capacidad del contenido generado por usuarios para generar interés sobre un tópico superando los contenidos generados comercialmente (Bickart & Schindler, 2001); el efecto del boca a boca en la confianza, lealtad e intención de compra (Awad & Ragowsky, 2008; Ekran & Evans, 2016; Chen, Fay & Wang, 2011; Pavlou & Ba, 2002) y recompra (Gauri, Bhatnagar & Rao, 2008). La gente común confía en las opiniones desinteresadas publicadas online, lo que sugiere que las empresas deberían enfocarse en los mecanismos que faciliten la dispersión del boca a boca (Duan, Gu & Whinston, 2008).

Influyentes y líderes de opinión

Dentro de las redes sociales surgen naturalmente figuras que tienen el potencial de influenciar un comportamiento de compra en sus contactos.

Un trabajo seminal, el modelo de comunicación de flujo de dos pasos, postula que las personas siguen a los líderes de opinión quienes a su vez son influenciados por los medios de comunicación (Katz & Lazarsfeld, 1955).

Desde ese trabajo a la actualidad el fenómeno de la influencia fue abordado desde diferentes enfoques y disciplinas. Los términos influyentes, influenciadores, líderes de opinión, *market mavens*, *hubs* o *alfa users*, se utilizan a menudo indistintamente para referirse a individuos que comparten un mismo conjunto de características.

Las empresas deben aprovechar la fuerza de la comunicación de consumidor a consumidor, atrayendo a grupos de usuarios que conecten con la marca y actúen posteriormente a su favor. Este grupo no necesariamente debe ser grande, pero sí influyente (Peters, Chen, Kaplan, Ognibeni & Pauwels, 2013; Risselada, Verhoef & Bijmolt, 2014).

La utilidad de identificar actores con rasgos de liderazgo de opinión dentro de una red social compleja es indudable para el marketing y abarca:

- Investigación de mercados.
- Muestreo y prueba de productos.
- Publicidad directa (Hawkins, Best, Coney & Carey, 1995).

A estos se pueden agregar:

- Relaciones públicas y eventos, involucrando a estos actores como motor de generación de *word of mouth* positivo, manteniendo bajo control el gasto en estas acciones con el máximo retorno sobre la inversión posible.
- Control de daños, operando sobre los actores más influyentes con el objeto de moderar y minimizar los daños a la imagen de la marca en caso de situaciones con ese potencial.

El accionamiento sobre individuos seleccionados mediante algún criterio particular para lograr un efecto multiplicador se conoce como sembrado o *seeding* y se da a través de la interacción de dos mecanismos: expansión del mercado y aceleración del consumo (Libai, Muller & Peres, 2013). El concepto de *seeding* ha sido abordado en varios trabajos. Como contribuciones que estudian el problema de optimización desde una perspectiva experimental se pueden mencionar a Chen, Wang & Wang (2010); Aral, Muchnik & Sundararajan, (2013); Kempe, Kleinberg & Tardos (2015) y Aghdam & Navimipou (2016). Para determinadas redes se han elaborado métricas derivadas de atributos y actividades propias de los usuarios para estimar, entre otros, indicadores de popularidad e influencia (Grossek & Holotescu, 2009).

Análisis de redes sociales

Las redes sociales online son en sí mismas entidades sociales que funcionan como un agregado de las conductas de sus componentes indivisos. Las redes son analizables principalmente en cuanto a su morfología (Smith, Rainie, Himelboim & Shneiderman, 2014), y a su vez cada uno de los integrantes de la red puede analizarse a partir de métricas que describen su posición dentro de la estructura de relaciones de la red (Hansen, Schneiderman & Smith, 2011).

Estudios previos sugieren que el análisis de la situación estructural de un actor dentro de una red es un buen indicador de liderazgo de opinión (Van der Merwe & van Heerden, 2009).

En *Marketing and social networks: a criterion for detecting opinion leaders* (Litterio, Nantes, Larrosa & Gómez, 2017) se propone una metodología para detectar actores influyentes dentro de redes sociales basado en la combinación de centralidad de intermediación y de autovector para clasificar los actores de una red social en cuanto a su nivel de influencia tal se observa en figura 1.



Figura 1. Matriz para detectar potenciales influenciadores dentro de una red social (Litterio et al, 2017).

Esta matriz hace uso de dos medidas globales de la importancia de un nodo en relación al resto de la red:

- Centralidad de autovector o *eigenvector*, es una medida estándar de influencia en una red que pondera las conexiones y la calidad de las conexiones de cada individuo.
- Centralidad de intermediación muestra en qué medida un actor se encuentra en el camino más corto entre otros dos nodos. Puede ser pensado como un “puente” dentro de la red, en que el actor toma una posición de intermediario o *broker*.

La matriz permite mapear los individuos que componen una red y de acuerdo con un criterio relevante para determinar los cortes, clasificarlos en cuatro segmentos con características de difusión diferenciadas.

Los líderes no necesariamente deben estar en el medio de cada red importante ya que esto sería a costa de una posición marginal en otra red. Existe un trade off en la construcción de este capital social (Balkundi & Kilduff, 2006).

El análisis de texto y sentimiento

El análisis de sentimiento (también conocido como minería de opinión) se refiere al uso de procesamiento de lenguaje natural, análisis de texto y lingüística computacional para identificar y extraer información subjetiva de unos recursos, con el objetivo de determinar la actitud de un interlocutor o un escritor con respecto a algún tema o la polaridad contextual general de un documento (Bodendorf y otro, 2009).

NodeXL permite hacer análisis de este tipo a partir del relevamiento de frecuencia de palabras sobre hasta tres grupos de palabras previamente definidos (Minqing y otro, 2004). Los grupos de palabras que se utilizan para hacer este tipo de análisis se denominan “léxicos”. Se utilizaron en este caso de estudio dos grupos de palabras adaptados de un léxico en idioma español que contiene palabras indicadoras de “agrado” vs. “desagrado” (Gravano y otro, 2014) con el objeto de determinar si existe polarización hacia uno u otro extremo en los comentarios de los actores detectados como influyentes, en relación a los que no lo son (Serrano Puche, 2016).

¿Qué es un hashtag?

El hashtag es una palabra o conjunto de palabras concatenadas precedidas por un símbolo numeral (#) que sirve para conectar el contenido de las redes sociales con un tema, evento, tema o conversación específicos. Los hashtags son ampliamente utilizados por las redes sociales online, donde funcionan como una etiqueta de metadatos facilitando a sistemas y usuarios el relacionamiento y agregación de contenidos sobre temas específicos.

Incluir hashtags en una publicación significa participar y ser visible en una conversación. Esto puede conducir a una mayor participación a través de 'me gusta', comentarios, reenvíos, acciones y nuevos seguidores. El origen del uso de hashtags en el contexto de redes sociales puede remitirse a una publicación en la red social Twitter del usuario Chris Messina (2007), en la que propuso la herramienta para destacar temas en esa red social.

Los hashtags son muy importantes como herramientas del marketing digital porque:

- Dirigen la comunicación a la audiencia adecuada
- Aumentan la probabilidad de generar interacción y conversiones
- Permiten dar contexto a un mensaje y expandir el alcance de la marca.

3- OBJETIVO

Este nuevo trabajo busca continuar la misma línea de investigación, centrándose ya no en la detección del influyente o de la forma de la red social (todos temas que ya fueron abordados anteriormente y que forman parte del análisis de la red social), sino en el contenido mismo del diálogo social.

Mediante un análisis ampliado de la red social que permite la detección de clusters a partir del contenido del cuerpo mismo de los mensajes intercambiados (mediante técnicas ya conocidas como la minería de texto) y con los influenciadores como foco, se propone un método que permite generar y ponderar mapas de audiencia, de los principales actores, y de los temas y sentimientos que predominan en cada subgrupo.

Con esta herramienta se busca complementar la visión del análisis de la red social como objeto de práctica de marketing, ya que abordamos en una primera instancia el método de detección de influyentes, en una segunda instancia la detección de clusters y hashtags, y por último le estamos aportando una información cualitativa (semántica) valiosa para la generación de acciones de marketing y de contenidos.

4- METODOLOGÍA

Propuesta de un método de trabajo

Se proponen un conjunto de pasos ordenados para lograr el objetivo deseado. Como supuesto previo se asume que existe una definición clara de la razón por la cual se desea analizar la red. Generalmente en marketing digital se habla del concepto de "conversiones" como la cantidad de casos en los que en una campaña o acción puntual se logra un efecto deseado.

Se trata definir, entonces, qué tipo de conversiones se quieren lograr: tráfico a un sitio web, generar conocimiento de marca, generar *leads* de venta, instalar o moderar el dialogo en torno a un tema en particular, etc.

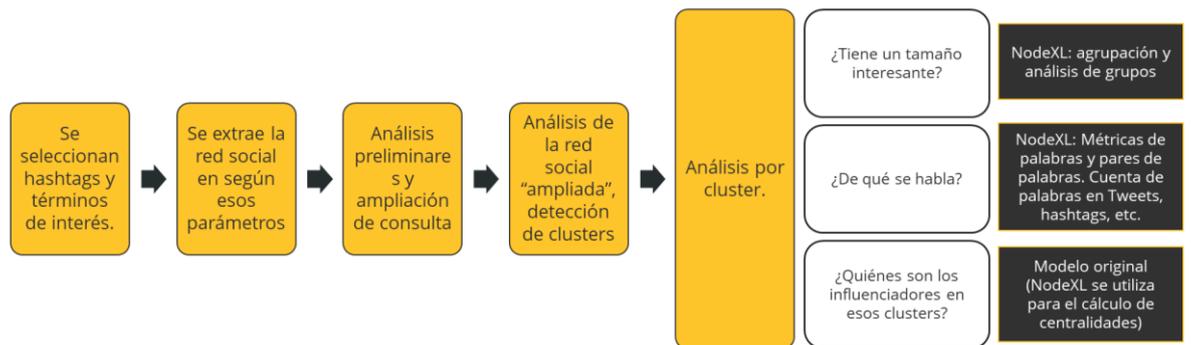


Figura 2. Proceso de análisis de redes sociales (Elaboración propia).

Tal como se muestra en figura 2, el primer paso del método propuesto consiste en seleccionar los hashtags y términos de interés relacionados con el objeto de análisis.

El segundo paso consiste en la extracción de datos de la red social en sí. En este trabajo se propone y se utiliza la aplicación NodeXL Pro, versión 1.0.1.445 (Smith et al., 2010). NodeXL es un paquete de *software* para Excel, de código abierto, para análisis y visualización de redes que incluye la posibilidad de importar datos directamente de las principales redes sociales (Hansen et al., 2011).

Cabe aclarar que este paso se puede realizar con varias herramientas diferentes, incluyendo lenguajes de programación como Python que disponen de módulos gratuitos para tal fin. Es importante aclarar también que las extracciones se realizan a través de la Interfaz de Programación de Aplicaciones (API) que ofrece cada red social, a través de las cuales moderan el tipo, cantidad, calidad y profundidad de datos que se pueden extraer. Esto a su vez cambia frecuentemente producto de actualizaciones en políticas de privacidad de datos.

Los tres primeros pasos deben concebirse como un proceso iterativo, ya que frecuentemente será necesario ampliar la consulta agregando o quitando descriptores de interés (hashtags, léxicos, descriptores geográficos y temporales, entre otros), en función del objetivo y alcance del análisis.

El producto será una red social ampliada sobre la que se realizará un análisis de detección de clusters. En este trabajo se propone el uso del algoritmo jerárquico propuesto por Clauset, Moore y Newman (2004), el cual se encuentra implementado en NodeXL.

Una vez detectados los clusters, se procede a analizar y ponderar cada uno de ellos a la luz del objeto de análisis. Se proponen tres instancias de análisis que buscan responder a preguntas concretas:

- **¿Tiene un tamaño y forma interesante?** Se analiza cada cluster en términos de su tamaño (cantidad total de actores) y de la densidad de conexiones (coeficiente de agrupamiento), que es un buen indicador de la calidad de relaciones y del diálogo

social dentro de esa red. Los clusters interesantes para el marketing serán generalmente aquellos más grandes y con mayor agrupamiento.

- **¿De qué se habla?** Se procede en segunda instancia con un análisis del texto en los *clusters* seleccionados en la instancia anterior. Esto tiene doble importancia: primero, sirve para lograr entender de qué se está hablando en cada *cluster* y el sentimiento prevalente, lo que permite enfocarse únicamente en grupos que sean de interés. Por otra parte, provee un listado de los términos que se utilizan en ese *cluster* y, fundamentalmente, los hashtags que permiten introducirse efectivamente en el diálogo.
- **¿Quiénes son los influenciadores?** Por último, se analizan los influenciadores de cada *cluster* según el modelo propuesto por Litterio et al, el cual permite detectar aquellos actores dentro de cada red o *cluster* que tienen rasgos de influencia.

Desarrollo de un caso

Como complemento a la descripción realizada sobre el método se demuestra su aplicación sobre un caso testigo. A diferencia de lo que se propone en párrafos anteriores, el caso testigo se desarrolla con una perspectiva exploratoria ya que el único objetivo es mostrar la factibilidad del método.

Se analiza una red social rica en contenido de texto (<https://www.twitter.com>) en la que se realizarán extracciones relacionadas con la empresa Mercado Libre (<https://www.mercadolibre.com.ar/>).

Extracción de la red y análisis preliminares

La extracción de la red se realiza mediante la herramienta de importación de Twitter Search Network de NodeXL en agosto de 2021 según la parametrización que se muestra en figura 3. Como se ve en la imagen, la consulta se realizó simplemente para tres términos clave (Mercado Libre escrito de diferentes formas comunes). La consulta o query a través de NodeXL admite programación mediante operadores avanzados para refinar la búsqueda y lograr mejores resultados.

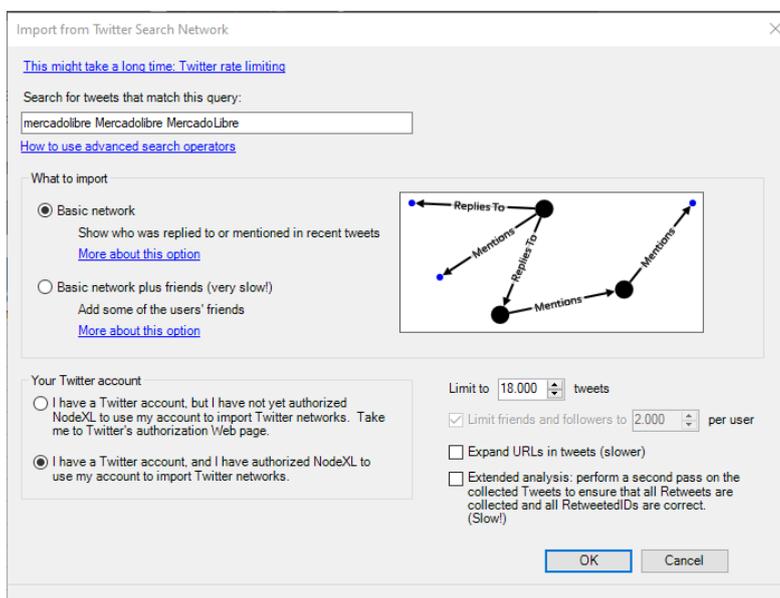


Figura 3. Opciones de importación de red desde NodXL.

La red obtenida mediante la consulta mostrada en figura 4 consta de 8.238 vértices o perfiles de usuario, los cuales se encuentran interconectados por un total de 21.171 edges o enlaces, que en el caso de esta red social en particular definen comentarios y reenvíos o retweets.

Graph Metric	Value
Graph Type	Directed
Vertices	8238
Unique Edges	9663
Edges With Duplicates	11508
Total Edges	21171
Self-Loops	2889
Reciprocated Vertex Pair Ratio	0,051326748
Reciprocated Edge Ratio	0,097641857
Connected Components	1709
Single-Vertex Connected Components	862
Maximum Vertices in a Connected Component	4639
Maximum Edges in a Connected Component	16133
Maximum Geodesic Distance (Diameter)	16
Average Geodesic Distance	4,040298
Graph Density	0,000159985
Modularity	0,454239
NodeXL Version	1.0.1.445

Figura 4. Importación de red desde NodeXL.

El grafo obtenido de la red, previo a cualquier tipo de procesamiento y de acuerdo a los parámetros que aplica NodeXL. Por defecto puede verse en figura 5, en la que cada punto corresponde a uno de los 8.238 vértices.

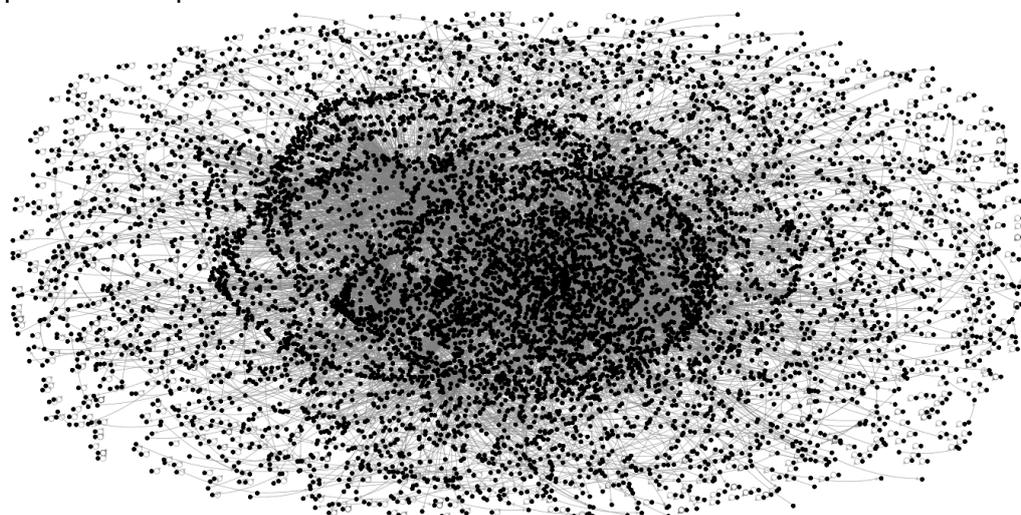


Figura 5. Grafo generado a partir de la extracción inicial (Elaboración propia con NodeXL).

Detección de clusters

Siguiendo con el proceso, se analizan los clusters presentes en la red social. Este proceso se encuentra automatizado en NodeXL a través de la opción *group by cluster*, tal como se muestra en figura 6.

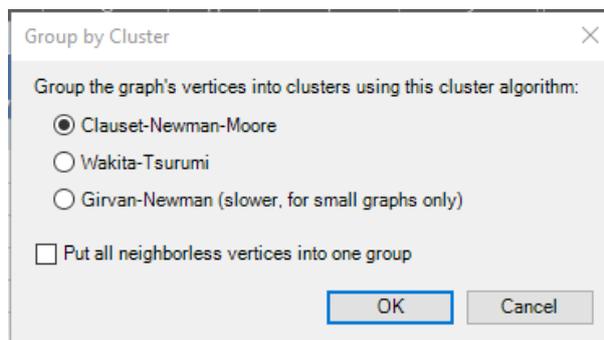


Figura 6. Menú de agrupación por cluster (NodeXL)

El proceso de detección de clusters es completamente automatizado. Una vez detectados los clusters se puede aplicar estilos y definir opciones de ordenamiento y visualización para a los diferentes grupos. En la figura 7 se puede ver una visualización de la misma red de la figura 5, pero en este caso se muestran cada uno de los clusters en una caja independiente, con su nombre, ordenados de mayor a menor cantidad de vértices y con estilos especiales definidos para los vértices de cada grupo.



Figura 7. Grafo de la red obtenida, clusterizada según el algoritmo de Clusset et al. en la misma figura de mapeó la centralidad de autovector al tamaño de los vértices (a mayor centralidad, mayor tamaño) y la centralidad de intermediación a la opacidad de los mismos (a mayor centralidad, mayor opacidad). (NodeXL)

El proceso de clusterización aplicado sobre la red anterior permitió la detección de 1.819 clusters, con un promedio de 4,53 vértices por cluster. Tal como puede verse en figura 7, la distribución de vértices por cluster no es homogénea, concentrando los 50 primeros clusters (el 3% de los grupos más grandes) el 57,6% de los vértices totales de la red.

Análisis de clusters: el tamaño y la forma

Los clusters que más interesan en marketing, tanto desde un enfoque descriptivo como positivo, son los de gran tamaño, por su alcance.

Ahora bien, esta métrica por sí sola no es suficiente. Es necesario analizar la forma de estos clusters, y una forma sencilla de hacerlo es a través del coeficiente de agrupamiento que indica la densidad de conexiones dentro del cluster, es decir el nivel de interacción entre las personas que forman ese cluster.

El cálculo del coeficiente de agrupamiento siempre da un resultado entre 0 y 1, indicando los valores pequeños bajo agrupamiento y los valores cercanos a 1 un gran nivel de interacción. Los clusters con altos niveles de interacción son interesantes para el marketing porque indican usuarios y comunidades activas, en las que se genera diálogo social. Los clusters con bajos niveles de interacción, generalmente corresponden a redes de *broadcasting*, en las que se emite un mensaje en forma unidireccional sin generar interacción entre los receptores, y se pueden identificar visualmente en el grafo porque tienen forma característica de estrella (un vértice central que conecta con varios vértices a su alrededor, mientras que ninguno de ellos conecta entre sí).

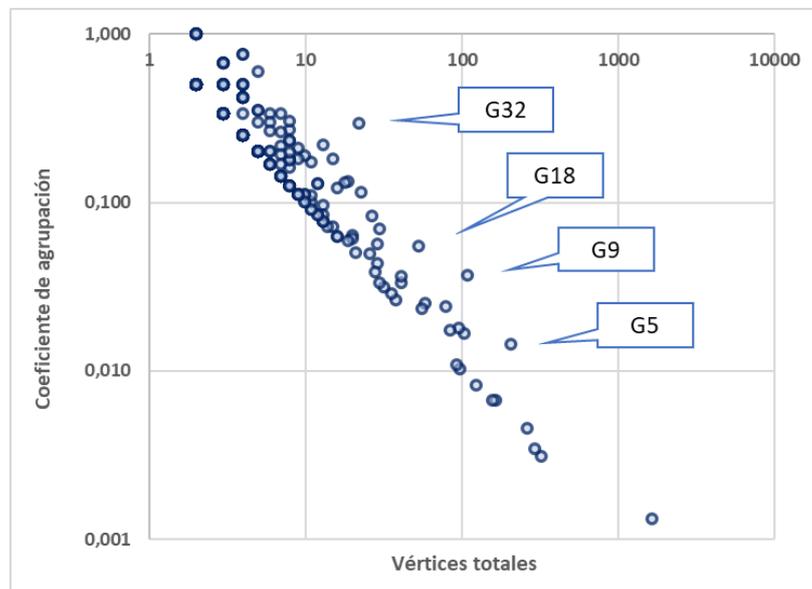


Figura 8. Diagrama de dispersión que muestra la cantidad de vértices por cluster (en eje X) contra el coeficiente de agrupación del cluster.

En la figura 8 puede verse un diagrama de dispersión en el que se grafican todos los clústeres obtenidos según su tamaño (en término de vértices totales) y de coeficiente de agrupación. Naturalmente existe una fuerte correlación entre ambas variables, ya que en grupos pequeños es más probable que se alcancen niveles altos de agrupación (tomando como extremo el grupo de una sola persona, cuyo coeficiente de agrupación por definición será igual a 1), sin embargo, es interesante ver como en esa misma figura pueden verse varios clusters que escapan la norma mostrando niveles altos de agrupación respecto a clusters de tamaño similar. Si se observan esos casos (G5, G9, G18, G32) en figura 7 puede verse a simple vista que presentan estructuras de

interconexiones complejas. Como contrapartida, clústeres como los G2 y G3 representan la forma de “estrella” que se mencionó anteriormente.

Análisis de clusters: texto

Se procedió a automatizar una parte importante del proceso de análisis de texto (como la tokenización) a través de NodeXL. También, se realizó un análisis de sentimiento (mediante un léxico en inglés que incorpora en todas sus versiones o mediante otros léxicos que se pueden ingresar según se ve en figura 9).

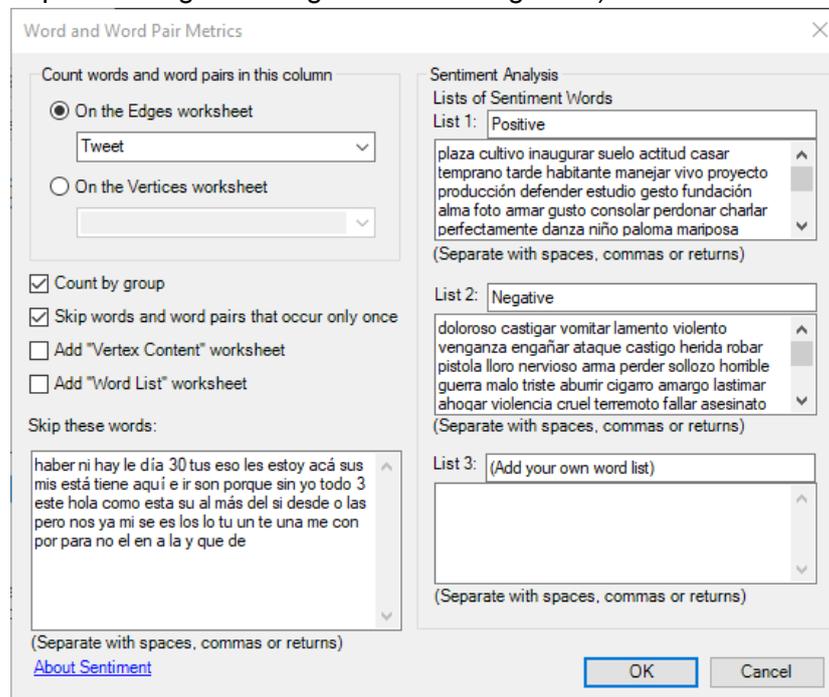


Figura 9. Herramienta para análisis de texto en NodeXL, donde se pueden ver el listado de palabras de valencia positiva y negativa, más el listado de palabras que no se deben ser tenidas en cuenta en el conteo.

Luego de correr ese proceso, se generaron dos tablas en las que se muestran los conteos totales de palabras y de pares de palabras. Para ambas tablas se muestran las métricas calculadas para toda la red y para cada cluster en particular, así como el conteo de palabras con valencia (tanto positiva como negativa). Esto permite identificar rápidamente de qué se habla en cada cluster, a su vez, y medir el sentimiento, permitiendo eventualmente integrar herramientas populares como las nubes de palabras.

De esas tablas se puede derivar un reporte como el que se muestra en figura 10, que permite visualizar en una única tabla los clusters detectados por NodeXL, el nivel de emocionalidad del diálogo en ese cluster (entendido como la sumatoria de las palabras con alguna carga afectiva implícita vs. total de palabras) y por último el sentimiento preponderante en ese cluster, a partir de las palabras que se utilizan.

Cluster	List #1	List #2	Tot Pala	Tamaño cluster	Coef agrupamien	EMOCIONALIDAD	SENTIMIENTO
G1	225	40	4316	1650	0,0013	6,1%	-1%
G2	6	0	65	327	0,0031	9,2%	14%
G4	4	1	63	263	0,0045	7,9%	-6%
G5	10	4	159	206	0,0143	8,8%	-15%
G6	18	1	221	165	0,0066	8,6%	8%
G7	33	7	486	157	0,0067	8,2%	-4%
G9	12	1	251	110	0,0369	5,2%	6%
G12	11	0	90	97	0,0177	12,2%	14%
G13	8	2	81	93	0,0109	12,3%	-6%
G14	11	1	164	85	0,0174	7,3%	5%
G15	15	1	358	80	0,0241	4,5%	7%
G18	9	0	147	53	0,0544	6,1%	14%

Figura 10. Reporte de sentimiento para una selección de clusters, elaboración propia.

Un análisis particular de cada cluster complementa este análisis y permite entender de qué se está hablando puntualmente en cada uno de esos clusters y cuáles son los hashtags que permiten introducirse en esa conversación. Así, podrían sacarse las siguientes conclusiones luego de un análisis posterior sobre algunos de los clusters que se muestran en figura 10:

Cluster	Palabras más frecuentes	Hashtags más frecuentes	
G5	eldestapeweb sms mensajes sabemos auspicia campaña	#mercadolibre #florenciorandazzo	Trata sobre una comunicación de campaña electoral de un candidato que supuestamente se hace desde teléfonos de contacto de Mercado Libre y se especula sobre la posibilidad de que estén vinculados o existe a una fuga de datos de usuarios de Mercado Libre.
G9	mercadolibre globant uala_arg tiendanube rastiarg despegar olxargentina	#unicornio #certifiedtechdeveloper	Se difunde y comenta la novedad de que Mercado Libre comienza a cotizar en la bolsa Argentina para finales de septiembre.
G18	mercadolibre orozco relaciones políticas públicas soniagarzagzz género	#mesadedebate #semanamujermipymehgo	Gira en torno de una charla con Aleheir Orozco (de Dir. de Políticas Públicas de Mercado Libre) y otras persona, sobre temas de género en el sector empresarial y otros temas afines (México).
G2	mercadolibre agradecerle empleados pintureria prestigio olivos vergüenza		Viralización de un tweet en que un usuario "agradece" a los empleados de una pinturería un destrato por una compra que realizó en Mercado Libre. Nótese en fig.10 que la polaridad del cluster es positiva, eso es problema del uso intenso del sarcasmo en el mensaje original y respuestas que no es correctamente interpretado por el análisis de sentimiento.
G4	descuento emocionada regresó pallete tuya	#miapalette #anmakeup1111	Anuncio y comentarios generados en torno del lanzamiento de un producto de cosmética "Mia Palette" en la plataforma de Mercado Libre México.
G12	libro nuestro propuestas rojoedwards mercadolibre comprarlo	#rutarepublicana	Anuncio y comentarios generados en torno del lanzamiento del libro "Ruta Republicana" en la plataforma de Mercado Libre Chile.

Análisis de clusters: influenciadores

De acuerdo al método propuesto por Litterio et al, se calculan las métricas de centralidad de autovector y de intermediación para cada uno de los vértices. Esto permite identificar los potenciales influenciadores dentro de la red en general. Las métricas calculadas para la red general son aplicables a cada uno de los clusters.

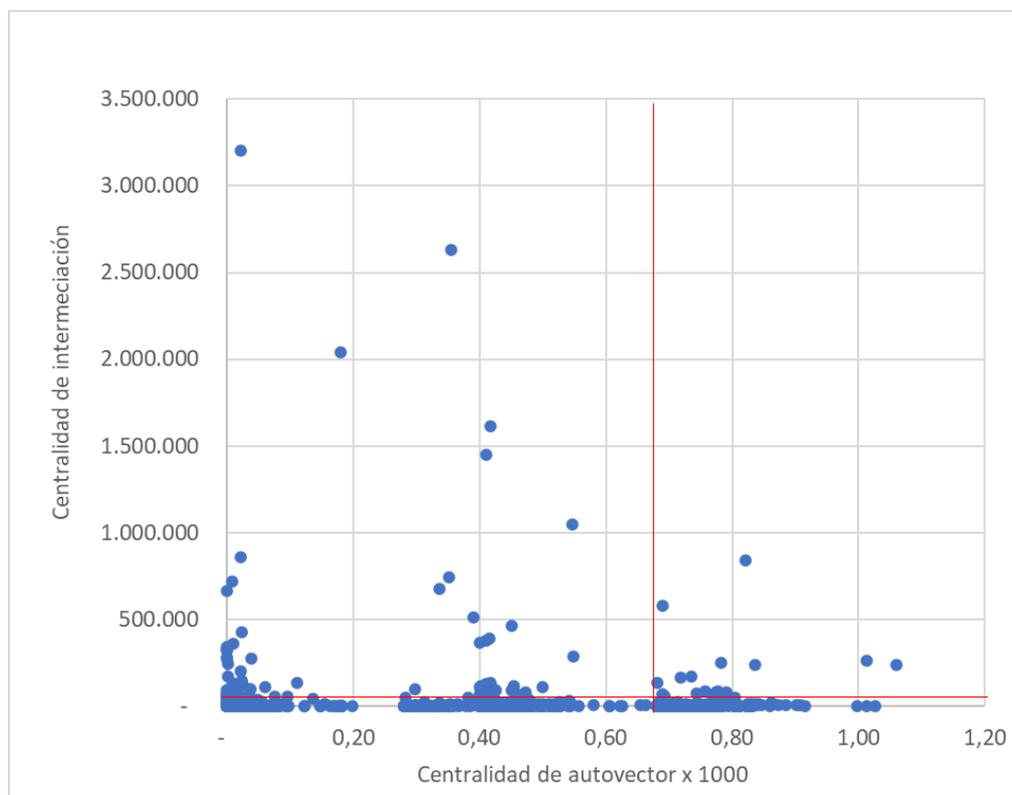


Figura 11. Diagrama de dispersión que muestra los vértices de la red mapeados según su centralidad de autovector (en eje X) y su centralidad de intermediación (en eje Y). Se excluyen las cuentas propietarias de Mercado Libre (ej mercadolibre, ml_ayuda, mercadopago).

Esto nos permite generar matrices de influencia con o la que se exhibe en la figura 11, donde se pueden ver todos los componentes de la red mapeados según la combinación de centralidades que sugieren Litterio et al. Además, el mismo modelo puede realizarse sobre los vértices que componen cada uno de los clusters detectados, permitiendo detectar quienes son los usuarios influyentes de cada cluster.

5- CONCLUSIONES

Se mostró la forma de analizar una red social y los grupos o clusters de individuos que componen esa red caracterizándolos a través de diferentes métricas, incluyendo el análisis semántico a partir de los léxicos que utilizan. Esto ayuda a ganar comprensión sobre el diálogo social en torno al tema que se elija, ponderar y elegir aquellos subgrupos de interés para maximizar la eficiencia del potencial mensaje para generar la conversión deseada, ya sea una venta, posicionar una idea, tráfico hacia un sitio web o red social, entre otros.

Por otra parte, se logró generar mapas de audiencia en términos de grupos que se comunican con respecto al tópico y tienen cierta cohesión. También cómo detectar a los influenciadores y extraer hashtags y keywords que ayudan a definir la audiencia y poder

“hablar su mismo idioma” de modo de mejorar la penetración del mensaje en el grupo objetivo y la posibilidad de generar diálogos de valor dentro de ese grupo.

La citada metodología se complementa con la propuesta de trabajos anteriores ya que tiene el potencial de lograr mayor eficacia y eficiencia en el marketing de redes sociales. Las palabras y hashtags más relevantes aseguran el alcance de una audiencia mayor.

6- BIBLIOGRAFÍA

Aghdam S. M. & Navimipou N. J. (2016) *Opinion leaders selection in the social networks based on trust relationships propagation*. Karbala International Journal of Modern Science 2, 2016 88-97.

DOI: 10.1016/j.kijoms.2016.02.002

Aral, S., Muchnik, L. & Sundararajan A. (2013) *Engineering social contagions: Optimal network seeding in the presence of homophily*. Network Science, 1, pp 125-153

DOI: 10.1017/nws.2013.6

Awad N. & Ragowsky A. (2008) *Establishing Trust in Electronic Commerce through Online Word of Mouth: An Examination across Genders*. Journal of Management Information Systems, Vol. 24, No. 4, pp. 101-121

Stable URL: <http://www.jstor.org/stable/40398913>

Bacile T., Ye C. & Swilley E. (2014) *From firm-controlled to consumer-contributed: consumer co-production of personal media marketing communication*. Journal of Interactive Marketing Volume 28, Issue 2, May 2014, Pages 117–133

DOI: 10.1016/j.intmar.2013.12.001

Balkundi, P. & Kilduff, M. (2006) *The ties that lead: A social approach to leadership*. The Leadership Quarterly 17 (2006) 419-439

DOI: 10.1016/j.leaqua.2005.09.004

Benedetti, A. (2015) *Marketing en redes sociales: Detrás de escena*. Buenos Aires: Ed AMDIA

Bickart B. & Schindler R. (2001) *Internet forums as influential sources of consumer information*. Journal of Interactive Marketing Volume 15, Issue 3 Pages 31–40

DOI: 10.1002/dir.1014

Bodendorf F., & Kaiser C. (2009) *Detecting Opinion Leaders and Trends in Online Social Networks*. SWSM'09, November 2, 2009, Hong Kong, China.

DOI: 10.1145/1651437.1651448

Chen W., C. Wang, & Y. Wang (2010) *Scalable Influence Maximization for Prevalent Viral Marketing in Large-Scale Social Networks*. Proceedings KDD'10.

DOI: 10.1145/1835804.1835934

Chen Y., Fay S. & Wang Q. (2011) *The Role of Marketing in Social Media: How Online Consumer Reviews Evolve*. Journal of Interactive Marketing Volume 25, pp 85–94

DOI: 10.1016/j.intmar.2011.01.003

- Clauset, A., Moore, C. & Newman, M. (2004) Finding community structure in very large networks. *Phys. Rev. E* 70, 066111 – Published 6 December 2004. DOI 10.1103/PhysRevE.70.066111
- Duan, W., Gu, B., & Whinston, A.B. (2008). *Do online reviews matter?—An empirical investigation of panel data*. *Decision Support Systems*, 45(3), 1007–1016
DOI: 10.1016/j.dss.2008.04.001
- Ekran I. & Evans C. (2016) *The influence of eWOM in social media on consumers' purchase intentions: An extended approach to information adoption*. *Computers in Human Behavior*. 61, pp 47-55.
10.1016/j.chb.2016.03.003
- Gauri D., Bhatnagar A., & Rao R. (2008) *Role of word of mouth in online store loyalty*. *Communications of the ACM* Vol 51 Issue 3 pp 89-91 .
DOI: 10.1145/1325555.1325572
- Gravano A. & Dell' Amerlina Ríos M. (2014) *Spanish DAL: A Spanish Dictionary of Affect in Language*. Reporte Técnico, Departamento de Computación, FCEyN-UBA.
- Hansen D., Shneiderman B., & Smith M. (2011). *Analyzing Social Media Networks with NodeXL: Insights from a Connected World*. Ed. Morgan Kauffman.
- Hawkins D., Best R., Coney K. & Carey K. (1995) *Consumer behavior: Implications for marketing strategy* McGraw-Hill
- Hewett K., Rand W., Rust R. & van Heerde H. (2016) *Brand Buzz in the Echoverse*. *Journal of Marketing*: May 2016, Vol. 80, No. 3, pp. 1-24.
DOI: 10.1509/jm.15.0033
- Hudson S., Huang L., Roth M. S. & Madden T. (2015) *The influence of social media interactions on consumer–brand relationships: A three-country study of brand perceptions and marketing behaviors*. *International Journal of Research in Marketing* V33
DOI: 10.1016/j.ijresmar.2015.06.004
- Katz, E. & Lazarsfeld, P.F. (1955) *Personal influence: The part played by people in the flow of mass communications*, New York: The Free Press.
- Kempe D., Kleinberg J. & Tardos E. (2015) *Maximizing the Spread of Influence through a Social Network*. *Theory of Computing Journal* Volume 11, Article 4 pp. 105-147.
DOI: 10.4086/toc.2015.v011a004
- King R., Racherla P. & Bush, V (2014) *What We Know and Don't Know About Online Word-of-Mouth: A Review and Synthesis of the Literature*, *Journal of Interactive Marketing* DOI: 10.1016/j.intmar.2014.02.001
- Lang, B. & Hyde K.(2013). *Word of mouth: what we know and what we have yet to learn*. *Journal of consumer satisfaction, dissatisfaction and complaining behavior* 26: 1–18.

- Libai B., Muller E. & Peres R. (2013) *Decomposing the Value of Word-of-Mouth Seeding Programs: Acceleration Versus Expansion*. Journal of Marketing Research: April 2013, Vol. 50, No. 2, pp. 161-176.
DOI: 10.1509/jmr.11.0305
- Litterio A. M., Nantes, E. A., Larrosa, J. M. & Gómez, L. J. (2017) *Marketing and social networks: a criterion for detecting opinion leaders*. European Journal of Management and Business Economics, Vol. 26 Issue: 3, pp.347-366.
DOI: 10.1108/EJMBE-10-2017-020
- Messina, Chris (2007) <https://twitter.com/chrismessina/status/223115412>
- Minqing H. & Bing L. (2004). *Mining and Summarizing Customer Reviews*. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining
DOI: 10.1145/1014052.1014073
- Nam H., & Kannan P. (2014) *The Informational Value of Social Tagging Networks*. Journal of Marketing: July 2014, Vol. 78, No. 4, pp. 21-40.
DOI: 10.1509/jm.12.0151
- Pavlou, P. & Ba, S. (2002) *Evidence of the Effect of Trust Building Technology in Electronic Markets: Price Premium and Buyer Behavior*. MIS Quarterly 26, 3, pp 243-268.
DOI: 10.2307/4132332
- Peters K., Chen Y., Kaplan A., Ognibeni B. & Pauwels K. (2013) *Social media metrics - A framework and guidelines for managing social media*. Journal of Interactive Marketing Vol 27 pp 281-298
DOI: 10.1016/j.intmar.2013.09.007
- Risselada H., Verhoef P., & Bijmolt T. (2014) *Dynamic Effects of Social Influence and Direct Marketing on the Adoption of High-Technology Products*. Journal of Marketing: March 2014, Vol. 78, No. 2, pp. 52-68.
DOI: 10.1509/jm.11.0592
- Serrano Puche, J. (2016) *Internet y emociones. Nuevas tendencias en un campo de investigación emergente*. Comunicar: Revista científica iberoamericana de comunicación y educación, Nº 46, pp 19-26.
DOI: 10.3916/C46-2016-02
- Smith, A.N., E. Fischer & C. Yongjian (2012) *How Does Brand-related User-generated Content Differ across YouTube, Facebook, and Twitter?*. Journal of Interactive Marketing 26. pp 102–113.
DOI: 10.1016/j.intmar.2012.01.002
- Smith, M., Milic-Frayling, N., Shneiderman, B., Mendes Rodrigues, E., Leskovec, J. & Dunne, C., (2010) *NodeXL: a free and open network overview, discovery and exploration add-in for Excel 2007/2010*, <http://nodexl.codeplex.com/> from the Social Media Research Foundation, <http://www.smrfoundation.org>

Smith, M., L. Rainie, I. Himelboim & B. Shneiderman (2014), *Mapping Twitter Topic Networks: From Polarized Crowds to Community Clusters*. Pew Research Center.

Van der Merwe R. & van Heerden G. (2009) *Finding and utilizing opinion leaders: Social networks and the power of relationships*. South African Journal of Business Management, Vol. 40, pp 65-76.

VanMeter R., Grisaffe D. & Chonko L. (2015) *Of “Likes” and “Pins”: The Effects of Consumers' Attachment to Social Media*. Journal of Interactive Marketing 32, pp 70-88.
DOI: 10.1016/j.intmar.2015.09.001

Wang Z. & Gon Kim H. (2017) *Can Social Media Marketing Improve Customer Relationship Capabilities and Firm Performance? Dynamic Capability Perspective*. Journal of Interactive Marketing. Volume 39, August 2017, Pages 15–26
DOI: 10.1016/j.intmar.2017.02.004