

Litterio, Mario; Nantes, Esteban A.

ANÁLISIS EN TIEMPO REAL DE LA INFLUENCIA EN REDES SOCIALES ONLINE: UN ENFOQUE AVANZADO

XXXVII EDUCA-AL

5 y 6 de Octubre 2023

Litterio, M, Nantes, E.A. (2023). Análisis en Tiempo Real de la Influencia en Redes Sociales Online: un enfoque avanzado. XXXVII Encuentro de Docentes Universitarios de Comercialización de Argentina y América Latina. Santa Fe, Argentina. En RIDCA. Disponible en:

<https://repositoriodigital.uns.edu.ar/handle/123456789/6502>



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/>



Análisis en Tiempo Real de la Influencia en Redes Sociales Online: Un Enfoque Avanzado

XXXVII 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

INDICE

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

1.- INTRODUCCIÓN

El presente trabajo se centra en el análisis de las redes sociales online y su relación con el marketing moderno.

Las redes sociales online permiten el diálogo entre individuos, marcas e instituciones, creando complejas redes de influencia social que afectan las opiniones y comportamientos de los consumidores.

Este trabajo da continuidad a investigaciones previas sobre la influencia en redes sociales online y su aplicación en el marketing digital. Se ha desarrollado una herramienta para analizar estas redes y detectar a los actores influyentes utilizando métricas de centralidad. Esta herramienta ha demostrado ser eficaz para identificar difusores de información en diferentes comunidades en línea (Litterio, Nantes, Larrosa & Gómez, 2017).

El análisis en tiempo real de los patrones de difusión de información en redes sociales tiene aplicaciones interesantes en los casos en que una respuesta o acción inmediata sobre la comunicación es valiosa. Modelos efectivos para detectar influyentes en estas redes y herramientas provenientes del aprendizaje automático (*machine learning*) como la detección automática de comunidades (clusterización) y el procesamiento automático de textos se convierten en instrumentos valiosos para gestionar información y tomar decisiones.

En la primera parte de este trabajo se desarrollarán conceptos teóricos que incluyen el marketing y el análisis de las redes sociales. Se introduce además información básica respecto a los métodos y técnicas utilizadas para obtener y procesar la información obtenida de Internet y a la problemática actual de X (antes Twitter) en particular.

En la segunda parte del trabajo, se desarrolla la metodología utilizada para el desarrollo del caso propuesto y se describen los resultados obtenidos.

Finalmente, se ofrecen las conclusiones y se discuten las limitaciones del trabajo, con el objetivo de promover futuras discusiones y extensiones de la investigación.

Es importante aclarar que Twitter desde mediados de 2023 se encuentra en proceso de migración de su nombre a X. Cuando en este trabajo se menciona a Twitter o a X se está hablando de lo mismo. De hecho la misma plataforma aún utiliza la palabra Twitter y derivados (por ejemplo *retweet*) extensivamente en su documentación y aplicaciones.

2.- MARCO TEÓRICO

El marketing y el boca a boca electrónico

La popularización de Internet y las redes sociales en particular cambiaron la forma en que las marcas se comunican con sus consumidores y potenciales consumidores. Anteriormente esa interacción era unidireccional, es decir de modo que la marca comunicaba un mensaje (publicitario) a una audiencia sin esperar que el receptor de ese mensaje respondiera.

Este paradigma se reflejó incluso en la primera iteración de Internet que hoy se identifica como Web 1.0 en la que las redes sociales aún no existían. En esos tiempos los sitios web de marcas que ya existían previo al surgimiento de internet (conocidas como *brick and mortar*) replicaban el formato de revista o catálogo de productos.

Dentro de ese paradigma, la interacción con el consumidor ha quedado reservada exclusivamente a la investigación e inteligencia de mercados (Bacile, Ye & Swilley, 2014; Benedetti, 2015).

El surgimiento y popularización de las redes sociales online en lo que se llama popularmente Web 2.0 generó un entorno en que los usuarios de las redes pueden comunicarse entre sí y con las marcas, facilitando la interacción y la creación de contenidos, opiniones y reacciones. Hewett y otros (Hewett, Rand, Rust & van Heerde, 2016) postularon 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.

Las redes sociales permitieron la existencia del boca a boca o *word of mouth*. El boca a boca electrónico o eWOM puede definirse conceptualmente como “cualquier declaración positiva o negativa realizada por clientes potenciales, actuales o anteriores sobre un producto o empresa, que se pone a disposición de una multitud de personas e instituciones a través de Internet”. Por lo tanto, eWOM ocurre cuando los consumidores generan e intercambian información o contenidos en línea en forma de opiniones sobre cualquier tema de su interés incluyendo publicaciones en redes sociales.

El eWOM ha sido extensivamente estudiado y reconocido como un factor influenciador 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).

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 actores individuales que tienen el potencial de influenciar la conducta 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).

La influencia en redes fue estudiada 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 la característica común de ser difusores eficientes de un mensaje o acción dentro de una red y generar realimentación positiva.

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 es útil en el marketing para investigación de mercados, muestreo y prueba de productos y publicidad directa (Hawkins, Best, Coney & Carey, 1995).

Más recientemente y en relación con las redes sociales online en particular pueden presentarse ventajas en la actividad de relaciones públicas ya que permite involucrar a estos actores para generar difusión positiva, manteniendo bajo control el gasto en estas acciones con el máximo retorno sobre la inversión. Además puede ser de utilidad para gestión de crisis en caso de eventualidades con potencial de daño a la imagen de marca para moderar y mitigar efectos negativos.

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 propios de los usuarios para estimar, entre otros, indicadores de popularidad e influencia (Grossek & Holotescu, 2009).

Análisis de redes sociales

El análisis de redes sociales es un campo de estudio interdisciplinario enfocado en el estudio de la interacción entre personas, cuya motivación inicial es el modelamiento de fenómenos sociales. Este análisis utiliza la teoría de grafos, una rama de la matemática que propone estructuras (grafos) compuestas por vértices (también llamados nodos o puntos) que están conectados por aristas (también llamadas enlaces o líneas) y permiten modelar relaciones entre objetos.

Las redes sociales online pueden considerarse en sí mismas como entidades sociales que funcionan como un agregado de las conductas de sus componentes. Se pueden describir y analizar con varias métricas a partir de su morfología (Smith, Rainie, Himelboim & Shneiderman, 2014), y a su vez por otro lado cada uno de los integrantes de la red puede analizarse y medirse individualmente a partir de métricas que describen su posición dentro de la estructura de relaciones de la red (Hansen, Schneiderman & Smith, 2011).

Encadenando con el tema anterior, diferentes estudios 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). 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* (compensación) en la construcción de este capital social (Balkundi & Kilduff, 2006).

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. 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.

El aprendizaje automático y el análisis de redes sociales online

Como en muchos otros campos del conocimiento, el análisis de redes sociales online se ha beneficiado de herramientas que permiten procesar información a través de

aprendizaje automático o *machine learning*. En el relevamiento presentado por Balaji, Annavarapu & Bablani (2021) se hace una enumeración actualizada de los diferentes algoritmos utilizados en el contexto del análisis de redes sociales online y casos de aplicación.

El aprendizaje automático forma parte de un conjunto de disciplinas que abordan el desafío de la inteligencia artificial desde la implementación computacional de métodos estadísticos y matemáticos. El objetivo es encontrar patrones y elaborar modelos que permitan hacer generalizaciones y predicciones sobre sistemas de grandes volúmenes de datos a través de algoritmos específicos. Existen diferentes problemas de aprendizaje a los que se puede aplicar aprendizaje de máquinas:

Aprendizaje supervisado: Utiliza datos clasificados o valorizados para producir un modelo predictivo que hace generalizaciones sobre diferentes estados del sistema y asocia a ellos un valor esperado de la variable objetivo. Se conoce como aprendizaje “supervisado” porque se le da al algoritmo un sistema completo de observaciones con las respuestas correctas para que “aprenda”. Se aplica comúnmente a problemas de clasificación y de regresión.

Aprendizaje no Supervisado: A diferencia del anterior, no se cuenta con una variable objetivo. El algoritmo busca relaciones en un cuerpo de datos y crea agrupaciones o conjuntos de reglas que los describan. Un caso típico de este tipo de aprendizaje es la detección automática de clusters o comunidades.

Aprendizaje por refuerzo o reinforcement learning: Es una forma de aprendizaje en la que no se dispone de un sistema de datos como en los dos casos anteriores, sino que se dispone de un entorno (habitualmente con infinitas posibilidades) en que una entidad o agente puede realizar acciones. El resultado de las acciones que tome determinará una recompensa que el agente debe buscar maximizar. Este enfoque es el que se utiliza en programación de vehículos que se autoconducen, juegos de computadora y aspiradoras automáticas entre otros.

El análisis de sentimiento aplicado a redes sociales

Existen múltiples ejemplos documentados y casos de uso exitosos del uso de información obtenida de redes sociales para, luego de las manipulaciones adecuadas, medir opinión pública respecto a un objeto de estudio.

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 & Kaiser, 2009).

Existen principalmente tres enfoques para el análisis de sentimientos. El primer enfoque es a través de aprendizaje automático, utilizando algoritmos y técnicas de aprendizaje supervisado y no supervisado. El segundo enfoque para el análisis de sentimientos es basado en léxicos. Estos métodos cuentan con diccionarios predefinido de palabras positivas y negativas junto con sus sinónimos, antónimos y su puntuación de polaridad. La puntuación y frecuencia de polaridad de las palabras positivas y negativas en una oración o documento permite inferir el sentimiento total de esa oración o documento. El tercer enfoque para el análisis de sentimientos es un enfoque híbrido que combina técnicas de aprendizaje automático y enfoques basados en léxicos (Chauhan, Sharma & Sikka 2021).

El análisis de sentimiento se está aplicando cada vez más en el ámbito del marketing y estudios sociales por su capacidad para procesar y clasificar rápidamente grandes volúmenes de datos. Un ejemplo es la aplicación al evaluar el estado de ánimo del público con el objetivo de prever los resultados de una elección en curso. La mayoría de

los estudios han pronosticado resultados electorales exitosos y positivos en sus trabajos de investigación, lo que demuestra el poder emergente de los datos de redes sociales y los enfoques computacionales inteligentes para realizar predicciones tan complejas (Chauhan, Sharma & Sikka 2021).

Existen modelos de predicción que monitorean desastres naturales basándose en el análisis de sentimientos de las entradas multimodales (Bryan-Smith et al. 2023).

Las redes sociales online como herramienta de formación de agenda periodística

Es conocido además y respaldado por estudios previos que redes como X tienen un fuerte impacto en la formación de agenda periodística.

Un estudio analizó la actividad de la audiencia en la construcción de la agenda periodística de cuatro medios líderes en el ámbito español y encontró que los medios presentan consistentemente portadas en las que hay un protagonismo más duradero para aquellos asuntos periodísticos que coinciden con los 'temas del momento' en Twitter. Los temas que los profesionales seleccionan reciben una atención periodística y un protagonismo más permanente cuando entran en sintonía con los asuntos que la participación ciudadana visibiliza mediante Twitter (Díaz, Langa & Medina, 2020).

En un estudio más reciente se establece que el uso de *tweets* por parte de los periodistas en las noticias es práctica estándar. El tratamiento de los *tweets* como fuente por parte de los periodistas sugiere la existencia de un bucle de retroalimentación que perpetúa la centralidad de Twitter en el ecosistema de noticias. A medida que Twitter se asocia con las noticias, los periodistas y formadores de opinión recurren a él para reaccionar a los eventos de noticias y recopilar información para a su vez utilizar como fuentes de información en sus historias (Molyneux & McGregor, 2022).

El análisis de redes sociales online en tiempo real

La definición de tiempo real es una convención subjetiva que depende del contexto específico en que se la utilice. Si bien se puede pensar en tiempo real como equivalente a inmediato o continuo, la implementación en la práctica de los sistemas que dependen del tiempo requieren generalmente de una actualización en intervalos discretos. La selección de la frecuencia en la que se actualice el sistema depende de las características y necesidades de la aplicación.

Como convención, en el contexto del análisis de redes sociales online podría definirse a un sistema que funciona en tiempo real a aquel que se actualiza con una resolución de 5 minutos. Una resolución mayor a la propuesta (actualizaciones más frecuentes) resulta poco práctico ya que no permite inferir cambios en una tendencia y por otro lado vuelve al sistema computacionalmente más demandante.

Existen en la actualidad varias alternativas para analizar las redes sociales online, pero pocas aún para hacerlo en tiempo real. Más aún, existen aplicaciones presentadas como tiempo real pero que se basan en cuerpos de datos históricos (la extracción de la información se realizó ex post y se utilizaron los *timestamps* o fechas de cada publicación para hacer el análisis longitudinal). Un ejemplo de este caso es el trabajo presentado por Khan y otros (2021).

El análisis en tiempo real de los patrones de difusión de información en redes sociales tiene aplicaciones interesantes en los casos en que una respuesta o acción inmediata sobre la comunicación es valiosa. Ejemplos de estos casos son en monitoreo de marcas durante lanzamientos de productos, el seguimiento de eventos masivos, el monitoreo de información respecto a opinión pública durante fechas de elecciones o el monitoreo de desastres naturales entre otros.

La reputación de una marca puede ser monitoreada en tiempo real y longitudinalmente, gestionada al aprovechar las relaciones recíprocas y virtuosas entre los impulsores, y conectada al desempeño financiero de la empresa (Rust et al. 2021).

Los beneficios indiscutibles de disponer de información en tiempo real en lugar de histórica son:

- Actualidad y precisión: El análisis en tiempo real otorga acceso a información actualizada y precisa sobre la dinámica de las redes sociales y la influencia de los actores clave. Esto permite tomar decisiones más informadas y oportunas.
- Identificación de tendencias emergentes: El análisis en tiempo real permite identificar tendencias emergentes, lo que es especialmente relevante en ámbitos donde la detección temprana puede brindar ventajas competitivas significativas.
- Respuesta ágil: El análisis en tiempo real permite monitorear y responder de manera ágil a los cambios en la opinión pública y las conversaciones en las redes sociales. Esto es valioso en situaciones de crisis, campañas políticas o lanzamiento de productos, donde una respuesta oportuna puede ser crucial.

X/Twitter: Características y novedades

Twitter o X es una red social online diseñada y creada originalmente para hacer publicaciones de texto cortas, que resultan fáciles de analizar. Tiene una gran cantidad de usuarios activos (350 millones en todo el mundo a agosto de 2023) y busca en general crear diálogo social facilitando el intercambio entre diferentes círculos sociales a diferencia de otras redes sociales que buscan compartimentalizar las interacciones dentro de comunidades más cerradas.

Twitter cuenta con una API o interfaz de programación de aplicaciones (<https://developer.twitter.com/>) que permite recopilar datos de la plataforma en forma masiva a través de consultas a enlaces específicos de la interfaz.

Este conjunto de características convirtieron a Twitter en una excelente fuente de datos para investigadores en opinión y ciencias sociales en general ya que ofrece una manera de observar reacciones reales de las personas a escala respecto a diferentes fenómenos. La información de Twitter se ha utilizado para fines tan diversos como realizar estudios de opinión política, difusión de información y noticias falsas.

La API de Twitter hasta febrero de 2023 fue gratuita y prácticamente sin límites para la cantidad máxima de datos que era posible extraer. A partir de esa fecha la plataforma desactivó el acceso gratuito y lo reemplazó por planes de pago con topes a las cantidades de datos descargables. El plan más económico en el momento que se publicó este artículo solo permite a terceros recopilar 10,000 datos por mes que es el 0,3% de lo que anteriormente se podía extraer con acceso gratuito en un solo día. Solo el plan más caro es equivalente en prestaciones al acceso anteriormente gratuito.

Esto ha hecho que sea mucho más difícil para los investigadores recopilar datos de Twitter, y ha llevado a la cancelación y abandono de múltiples estudios y aplicaciones que contaban con el acceso en todo el mundo.

No existe al momento otra plataforma que sustituya a Twitter en términos de cantidad de usuarios activos y calidad de la información.

3.- OBJETIVO

Los objetivos de este trabajo son:

- 1.- Desarrollar un marco teórico y metodológico que permita el análisis en tiempo real de la influencia en redes sociales online.
- 2.- Utilizar herramientas avanzadas de análisis de redes sociales y aprendizaje automático para interpretar y visualizar los mapas de audiencia e interacciones en tiempo real.

3.- Identificar y analizar grupos o clusters consolidados de individuos en tiempo real, con el objetivo de comprender su comportamiento y sus características.

4.- Identificar influenciadores en tiempo real y extraer palabras claves relevantes para definir y comprender la audiencia en tiempo real.

Con este trabajo se busca en definitiva lograr una visión integral del análisis de la red social desde el punto de vista del marketing, ya que se abordan varios temas como la detección de influyentes, la detección de clusters y por último información sobre el contenido del diálogo social que se convierten en elementos valiosos para la generación de acciones de marketing y de contenidos.

4.- METODOLOGÍA

Extracción de información desde X/Twitter

Se analizó contenido de la red social online Twitter con extracciones hechas desde la API de la plataforma a través de programación en lenguaje Python (<https://www.python.org>), con el objetivo de crear una red compuesta por usuarios de la red (nodos) y enlaces creados a partir de comunicación entre ellos (aristas).

Para realizar la extracción de las publicaciones se creó una pieza de código o *script* para recuperar información específicamente a través del endpoint de la API **Recent Search** (Search Tweets Introduction, n.d.) con una consulta o *query* específico para obtener las publicaciones que tenían en su contenido un objetivo de búsqueda.

Este *script* se programó de modo tal que se ejecutó automáticamente cada 5 minutos en la franja de tiempo que duró la “escucha” en tiempo real, acumulando los datos extraídos en una única base de referencia.

Cada uno de los registros de esa base contiene un *tweet* que puede ser una publicación original, un *reply* en los casos en que la publicación consiste en una respuesta a otro usuario o un *retweet* en caso que correspondiera a una replicación de un *tweet* preexistente además de datos de su autor. Estos registros son en concreto las aristas que formaron parte de la red social analizada.

Como para analizar la red de usuarios además de aristas se necesitan nodos -y en la base original solo se dispone del usuario que creó la publicación, se hizo necesario rastrear en la mayoría de los casos al segundo usuario, quien creó originalmente el *tweet* o al menos el *retweet* anteriormente.

La identidad del segundo usuario se rastreó a través del endpoint **Tweets lookup** (Tweets Lookup Introduction, n.d.), la cual luego se combinó con la base original para completar los perfiles de usuarios e identificarlos individualmente.

Procesamiento de los datos

Parte del análisis de redes sociales se realizó a través de NetworX (Hagberg, Schult & Swart, 2008). NetworkX es una librería abierta de Python para la creación, manipulación y estudio de la estructura, dinámica y funciones de redes complejas.

Para la visualización se utilizó además el software de código abierto para la visualización y análisis de redes Gephi (Bastian, Heymann & Jacomy, 2009) que resulta más versátil y potente para crear visualizaciones, incluyendo la posibilidad de crear visualizaciones dinámicas (que reflejan la evolución de la red en función del tiempo) y análisis adicionales.

Los datos se fueron monitoreando a medida que se extraían con lo cual se cumplió con la condición de análisis en tiempo real. Además se agregaron análisis post-hoc que permitieron entender con una visión retrospectiva el tipo de hallazgos que se pueden detectar con este tipo de técnicas.

Elección de un caso de estudio

Se optó por seleccionar un tema de relevancia política (por el alto grado de involucramiento o *engagement* que suelen tener estos temas en la plataforma) para analizar el funcionamiento y utilidad del monitoreo en tiempo real.

La extracción de información o período de escucha se realizó el día 7 de agosto de 2023 entre las 15:00 horas y 21:40 horas (hora local Argentina).

Se utilizó como clave de búsqueda el texto “ministro de economía”, que ya se encontraba convertido en tendencia y se consideró relevante para medir la sensibilidad del público ante una corrida cambiaria que se dió ese mismo día en el valor del Dólar Estadounidense relativo al Peso (moneda local). En concreto la unidad de dólar informal cotizó a pesos 594 a las 16:05 horas del 7 agosto, lo cual implicó una suba del 3,48% con respecto a los datos que tuvo durante la apertura de operaciones (Dólar Blue Hoy: A Cuánto Cerró La Cotización Este Lunes 7 Agosto, 2023).

Es importante aclarar que las consultas o *queries* que se presentan a la API para delimitar el contenido de interés pueden involucrar texto, hashtags y utilizar varios descriptores disponibles además de operadores lógicos.

Resultados generales

Los datos se fueron analizando a medida que se obtuvieron de la red. El período de escucha comprendido entre las 15:00 y 21:40 implicó 80 corridas del sistema de 5 minutos cada una con un promedio de 70 *tweets* obtenidos en cada intervalo.

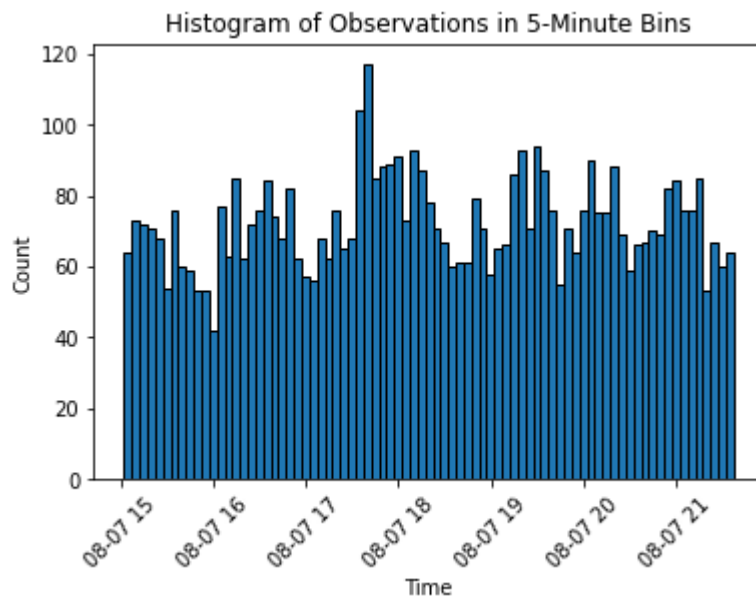


Fig. 1: Cantidad de tweets obtenidos para cada intervalo de tiempo durante el período de escucha.

Detección de comunidades

En la figura 2, se muestran visualizaciones de la red generadas en tiempo real por Gephi en diferentes momentos de la extracción. En la última figura se pueden ver 4.547 nodos totales más 5.631 aristas que componen la red obtenida acumulada a las 21:40 hs. La disposición de los nodos corresponde con el algoritmo ForceAtlas2 (Jacomy et al. 2014).

Los colores corresponden a diferentes agrupamientos calculados por el mismo software según un algoritmo de aprendizaje automático no supervisado que calcula la modularidad (Blondel, Guillaume, Lambiotte & Lefebvre, 2008).

La modularidad es una medida de la estructura de las redes que mide la fuerza de la división de una red en módulos (también llamados grupos, clusters o comunidades). En el contexto del presente trabajo, los grupos contienen los usuarios afines por replicar y responder mensajes dentro de un mismo círculo de usuarios.

El tamaño de cada nodo a su vez se correlaciona con la centralidad de vector propio calculada para cada usuario.

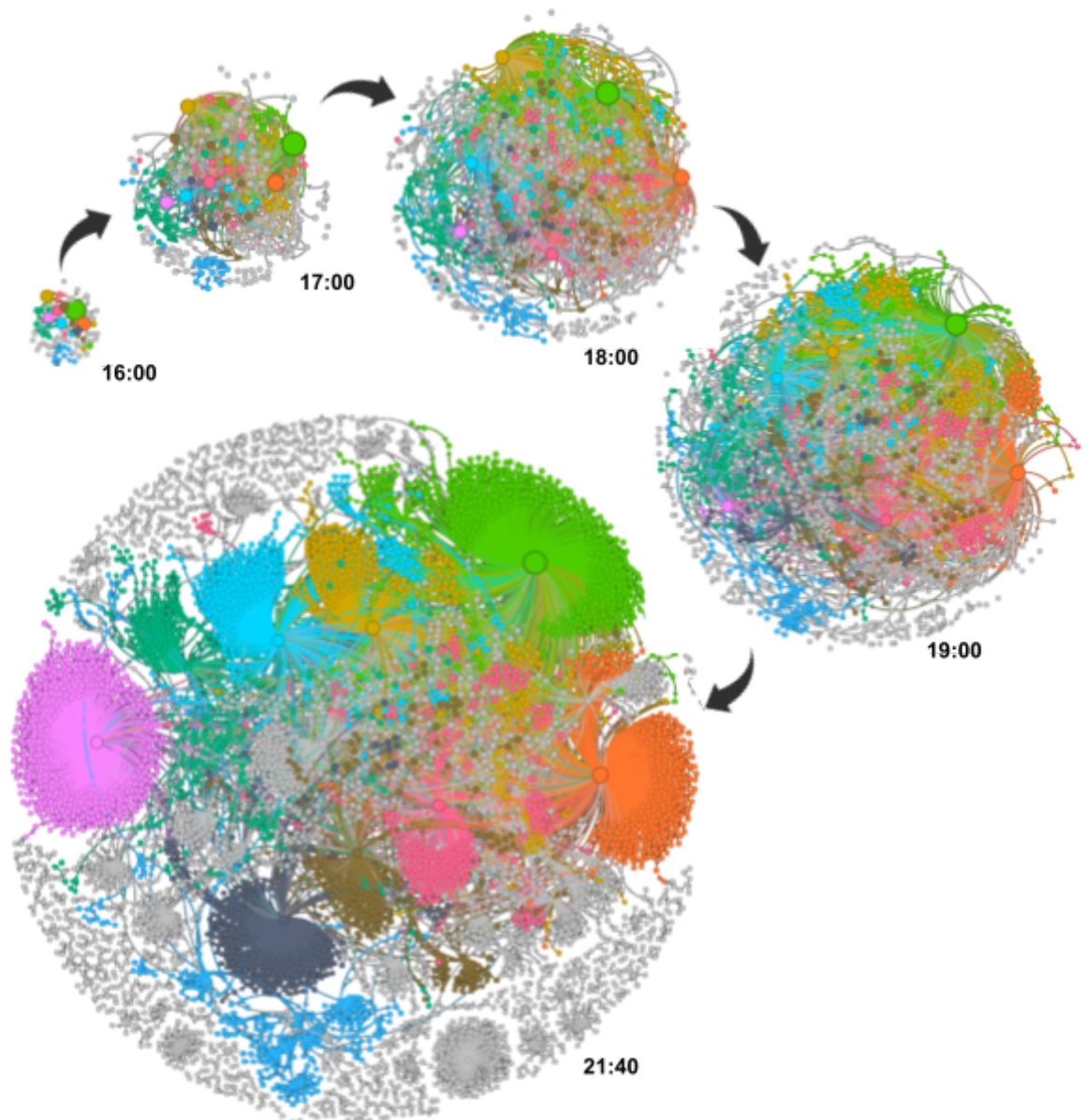


Fig. 2: Capturas de la red en diferentes intervalos de tiempo. Elaboración propia con Gephi. En la imagen puede verse cómo se pueden detectar en tiempo real diferentes comunidades de difusión, como van definiendo su audiencia e incluso algunas cambiando de posición en el tiempo, mientras otras se mantienen virtualmente alejadas de la conversación con sus propias audiencias.

El propósito de calcular clústeres es poder identificar subcomunidades dentro de la red obtenida. Esto ayuda a ganar comprensión sobre el diálogo social y crear mapas de

audiencia en torno al tema que se elija, ponderar y elegir aquellos subgrupos de interés para maximizar el potencial del mensaje en generar la conversión deseada, ya sea una venta, posicionar una idea, tráfico hacia un sitio web o red social, entre otros (Nantes & Litterio, 2021).

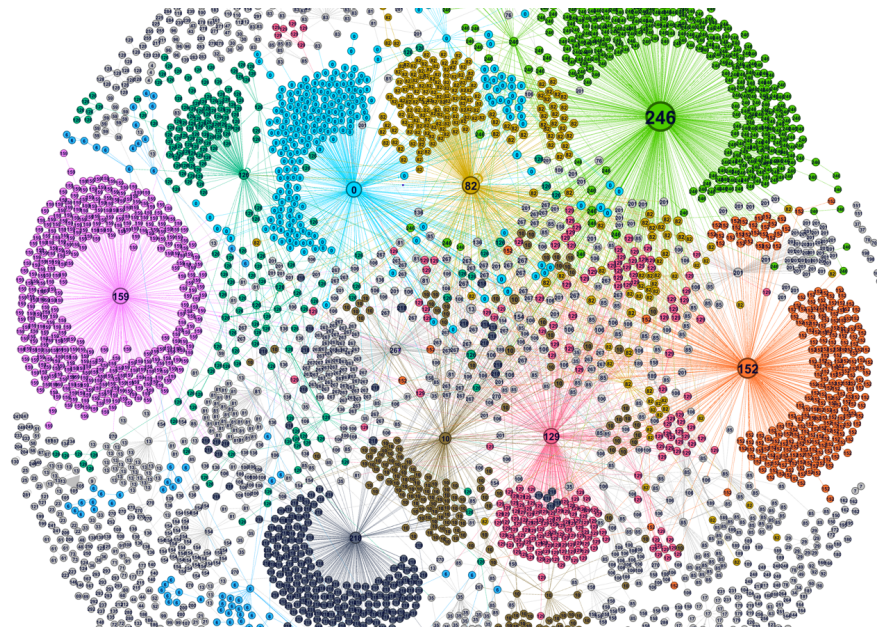


Fig. 3: Captura parcial de la red social con el número de módulo o cluster sobreimpreso sobre cada nodo.

Estos clusters se pueden analizar en cuanto a su tamaño, cercanía e interacciones, pero también longitudinalmente. En la figura 4 puede verse cómo evoluciona el tamaño de cada grupo a medida que transcurren los 80 intervalos de 5 minutos que se relevaron. La figura permite ver cómo se van introduciendo diferentes comunidades en la conversación.

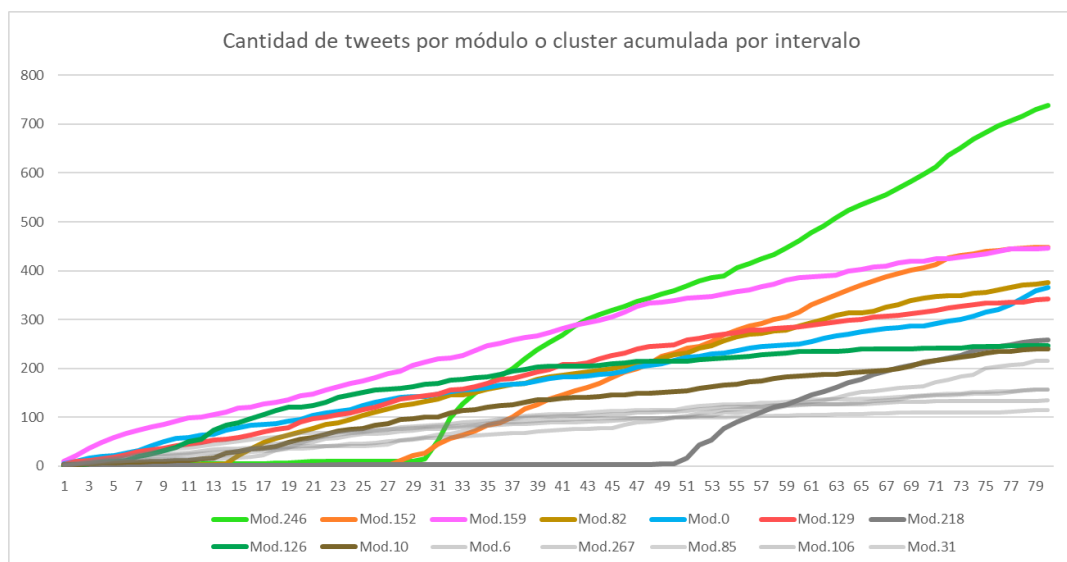


Fig. 4: Evolución de clusters en términos de tweets acumulados para los 14 principales clusters detectados. El color de cada línea replica los colores de figuras anteriores. Obsérvese cómo el módulo 246 (línea color verde claro) ingresa en la comunicación tardíamente, pero con pocos intervalos transcurridos se puede inferir su crecimiento futuro.

Detección de influyentes

El análisis de módulos o clusters permite detectar mapas de audiencia en términos de grupos que se comunican con respecto al tópico y tienen cierta cohesión. Adicional a esto, detectar a los influenciadores a través de sus medidas de centralidad dentro de la red y extraer hashtags y palabras o frases que ayudan a definir la audiencia y entender de qué habla cada grupo.

En tabla 1 pueden verse los módulos o clusters principales, el Tweet modal (más reenviado dentro de una misma comunidad sin modificaciones), el nombre del usuario influyente dentro de esa comunidad y su centralidad de vector propio y de intermediación.

Modulo	Tweet modal	Nombre de usuario influyente	Centralidad de vector propio	Centralidad de intermediación
0	RT @javierlanari: Hablando en serio. No podés ser ministro de Economía y candidato a Presidente y decir esta burrada. Y lo dice convencido....	javierlanari	0,3861	1.007.484
6	RT @julia13430: El dólar casi a 600 pesos por favor a donde vamos a ir a parar si llega a ser el señor ministro de economía presidente, pob...	diegobranca	0,0147	199.402
10	RT @maquialifraco: Un dato poco conocido es que Lousteau no renunció por elección a su cargo de Ministro de Economía de Cristina Kirchner....	maquialifraco	0,2144	410.763
82	RT @lauritalonso: Sergio Massa, el ministro de economía, quiere ser Presidente. ¿Cómo va a arreglar algo si empeoró todo en un año? https://...	lauritalonso	0,4713	805.771
85	RT @Daniel67735075: DOLAR 596 PESOS La caradurez al palo	Daniel67735075	0,0398	77.285
106	RT @JorRausch: 🇦🇷 #ARGENTINA AR EN UNA GRAVE CRISIS ECONÓMICA	Mimichas1959	0,0280	12.830
126	RT @edimburghok: Sergio Massa el peor ministro de economía de los últimos 40 años, el de las infulas presidenciales... canalla!!	edimburghok	0,1208	533.557
129	RT @loderaulo: Argentina país bizzarro, el único país donde un Ministro de Economía que va fundiendo al país sin plan alguno se presenta co...	loderaulo	0,3811	790.932
152	RT @LicTraviata: \$600 el dólar y No hay presidente	LicTraviata	0,5857	1.128.346
159	RT @therealbuni: Creo que salir en la página del dolar siendo ministro de economia no es la mejor propaganda pero como no soy candidato a p...	therealbuni	0,3981	1.429.088
246	RT @FerIglesias: Dólar a 600? Quédense tranquilos, que el ministro de Economía @SergioMassa se está ocupando del tema.	FerIglesias	1,0000	1.817.513
267	RT @vivaroca2015: No sé qué espera el presidente Massa para pedirle la renuncia al ministro de economía.	vivaroca2015	0,1658	318.141

Tab. 1: Caracterización de comunidades principales a través de sus tweets modales e influyentes.

Nótese cómo en algunos casos la comunidad se desarrolla en torno a una publicación original creada por el mismo influyente de ese módulo (por ejemplo en el módulo 246 el usuario influyente FerIglesias es el creador del tweet modal). En otros casos el influyente se encuentra fuera del diálogo y es el destinatario de los tweets (por ejemplo en el módulo 6 el usuario diegobranca es el receptor de varios mensajes sin necesariamente involucrarse con ninguno).

5- CONCLUSIONES

En este trabajo se propuso, delineó y ejecutó un método para relevar y analizar en tiempo real información de la red social X o Twitter.

Se demostró a través de un caso cómo la información obtenida con el método permite hacer un seguimiento en tiempo real del surgimiento, evolución y caracterización de grupos de afinidad en función de la difusión de contenidos.

Entendemos que metodologías como la propuesta pueden ser fuente de ventajas competitivas

Esta publicación se extenderá a futuro mediante la complementación con hallazgos de trabajos anteriores, entre ellos el análisis del diálogo social en torno al tema que se elija, ponderar y elegir aquellos subgrupos de interés para maximizar el potencial del mensaje en generar la conversión deseada, ya sea una venta, posicionar una idea, tráfico hacia un sitio web o red social, entre otros.

Una limitación presente en esta línea de investigación en general y este trabajo en particular tiene que ver con la disponibilidad de la información necesaria para desarrollar

modelos, probarlos y publicarlos. El acceso a los datos de las redes sociales se ha vuelto más restrictivo y las empresas que controlan los principales medios han tomado medidas que progresivamente vuelven más complejo e inaccesible el acceso a los datos.

Para el caso de Twitter en particular, esta limitación podría salvarse si la empresa habilita un acceso académico ilimitado a su API (el cual ha anunciado pero no implementado aún), o recurriendo a métodos alternativos como el data scraping.

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.
- Aral, S., Muchnik, L. & Sundararajan A. (2013) Engineering social contagions: Optimal network seeding in the presence of homophily. *Network Science*, 1, pp 125-153
- 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
- Balkundi, P. & Kilduff, M. (2006) The ties that lead: A social approach to leadership. *The Leadership Quarterly* 17 (2016) 419-439
- Bastian M., Heymann S., Jacomy M. (2009). Gephi: an open source software for exploring and manipulating networks. *International AAAI Conference on Weblogs and Social Media*.
- 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
- Balaji, T. K., Annavarapu, C. S. R., & Bablani, A. (2021). Machine learning algorithms for social media analysis: A survey. *Computer Science Review*, 40, 100395.
- Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008(10), P10008.
- Bodendorf F., & Kaiser C. (2009) Detecting Opinion Leaders and Trends in Online Social Networks. *SWSM'09*, November 2, 2009, Hong Kong, China.
- Bryan-Smith, L., Godsall, J., George, F., Egode, K., Dethlefs, N., & Parsons, D. (2023). Real-time social media sentiment analysis for rapid impact assessment of floods. *Computers & Geosciences*, 178, 105405.
- Chauhan, P., Sharma, N., & Sikka, G. (2021). The emergence of social media data and sentiment analysis in election prediction. *Journal of Ambient Intelligence and Humanized Computing*, 12, 2601-2627.
- Chen W., C. Wang, & Y. Wang (2010) Scalable Influence Maximization for Prevalent Viral Marketing in Large-Scale Social Networks. *Proceedings KDD'10*.
- 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
- Clauset, A., Moore, C. & Newman, M. (2004) Finding community structure in very large networks. *Phys. Rev. E* 70, 066111 – Published 6 December 2004. DOI
- Díaz, P. L. P., Langa, E. A., & Medina, R. Z. (2020). La construcción de la agenda de los cibermedios. Estudio comparativo con las preferencias temáticas de lectores y usuarios de Twitter. *Revista Latina de Comunicación Social*, (75), 225-244.
- Dólar blue hoy: a cuánto cerró la cotización este lunes 7 agosto. (2023, Agosto 7). - El Cronista.
<https://www.cronista.com/finanzas-mercados/dolar-blue-hoy-a-cuanto-cerro-la-cotizacion-este-lunes-7-agosto/>

- 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
- 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.
- 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.
- Hagberg Aric A., Schult Daniel A. and Swart Pieter J.(2008) Exploring network structure, dynamics, and function using NetworkX, in *Proceedings of the 7th Python in Science Conference (SciPy2008)*, Gael Varoquaux, Travis Vaught, and Jarrod Millman (Eds), (Pasadena, CA USA), pp. 11–15, Aug 2008
- 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.
- 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*
- Jacomy M, Venturini T, Heymann S, Bastian M (2014) ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software. *PLoS ONE 9(6)*: e98679.
- 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.
- Khan, H. U., Nasir, S., Nasim, K., Shabbir, D., & Mahmood, A. (2021). Twitter trends: a ranking algorithm analysis on real time data. *Expert Systems with Applications*, 164, 113990.
- 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*
- 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.
- 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.
- Molyneux, L., & McGregor, S. C. (2022). Legitimizing a platform: Evidence of journalists' role in transferring authority to Twitter. *Information, Communication & Society*, 25(11), 1577-1595.
- Nam H., & Kannan P. (2014) The Informational Value of Social Tagging Networks. *Journal of Marketing*: July 2014, Vol. 78, No. 4, pp. 21-40.
- Nantes E. & Litterio, A.M. (2021) *Redes sociales online: Análisis de textos para generar contenidos para distintos segmentos*. XXXV EDUCA-AL, Santiago del Estero, Argentina.

- 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.
- 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
- 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.
- Rust, R. T., Rand, W., Huang, M. H., Stephen, A. T., Brooks, G., & Chabuk, T. (2021). Real-time brand reputation tracking using social media. *Journal of Marketing*, 85(4), 21-43.
- Search Tweets introduction. (n.d.). Search Tweets Introduction | Docs | Twitter Developer Platform. <https://developer.twitter.com/en/docs/twitter-api/tweets/search/introduction>
- 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.
- Smith, M., L. Rainie, I. Himelboim & B. Shneiderman (2014), Mapping Twitter Topic Networks: From Polarized Crowds to Community Clusters. Pew Research Center.
- Tweets lookup introduction. (n.d.). Tweets Lookup Introduction | Docs | Twitter Developer Platform. <https://developer.twitter.com/en/docs/twitter-api/tweets/lookup/introduction>
- 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.
- 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