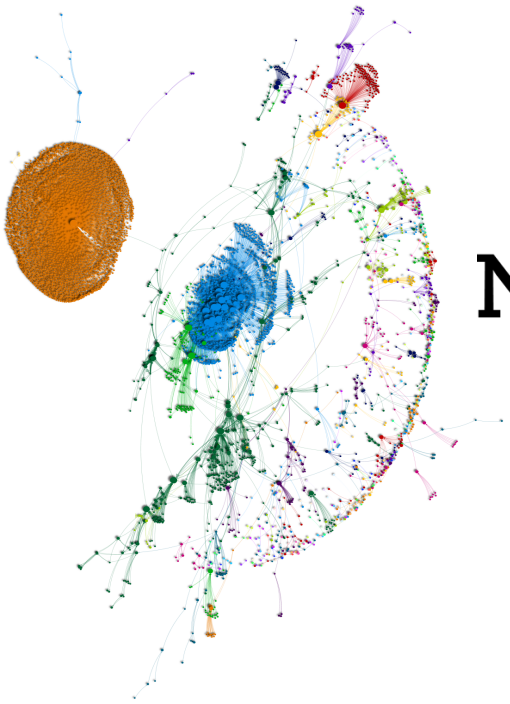


NodeXL como puente entre el Big Data y Thick Data para el análisis netnográfico de redes sociales.

Autor: Marcos De Colsa

Publicado: Junio-01-2023

A project from the socialmedia RESEARCH FOUNDATION

Introducción

Este artículo nace con la idea de escribir piezas de difusión para **NodeXL**, en el marco de las pláticas e interacciones que he tenido con los miembros de la **Social Media Research Foundation**, con el objetivo de difundir esta herramienta de investigación entre el público de habla hispana.

El mundo de la investigación digital, en todas sus vertientes, cuantitativa y cualitativa es profundamente desigual entre el público de habla inglesa y el público de habla hispana, la gran mayoría de los recursos, artículos y libros se encuentran en inglés, y sobre todo la literatura más actual en donde se pueden leer los avances en el terreno de las metodologías de investigación y las herramientas tecnológicas que se utilizan, está lejos de ser traducida al español en un lapso de tiempo pertinente antes de que la información quede obsoleta.

Todos los temas que se tocan en este artículo son abundantes en la bibliografía anglosajona, ya sea por separado o en trabajos que relacionan algunas de las temáticas, sin embargo, para lector de habla hispana es difícil encontrar lecturas en español que crucen aspectos que hilan cierta narrativa como en este caso la Revolución Digital, el Análisis de Redes Sociales y herramientas de trabajo como **NodeXL**.

Hay una relación no tan obvia entre todos estos temas, pero quien se interese en entrar al mundo de la investigación utilizando **NodeXL**, o quienes deseen tener una aproximación sólo teórica, tendrán una mejor y más amplia perspectiva de cuál es la lógica metodológica y de investigación y el alcance que se puede tener con esta herramienta. Es decir, entender el contexto en el cual hacemos investigación también es un medio y un abanico de posibilidades que nos permite ampliar nuestro panorama en cómo utilizamos nuestras herramientas de trabajo.

Entonces en este texto hablaremos de Big Data, Thick Data, Análisis de Redes Sociales y Netnografía, además de **NodeXL**, sin pretender que sea un texto intensivo sobre todos estos temas, ya que sería imposible en un escrito corto de divulgación, pero lo que sí pretende este trabajo es ser una provocación para que los lectores se interesen por los diversos temas y los agoten por su cuenta, ya sea en conjunto o por

separado, y también es una invitación para quienes les sea posible desarrollar material en español y tratar de reducir, en la medida de lo posible, la brecha de acceso al conocimiento entre el mundo hispano y anglosajón.

Detrás de este texto que se presenta, también hay mucho trabajo en la práctica profesional, la Dra. Verónica Espinoza -también integrante del equipo de la **SMRF** y yo hemos utilizado **NodeXL**, desde el 2012 (hace diez años), en ese momento **NodeXL** nos servía para bajar datos de ciertas marcas y sus seguidores en Twitter, es decir, nuestro interés se encontraba en identificar las comunidades digitales de diferentes marcas con las que trabajábamos y **NodeXL** era -y sigue siendo- la herramienta perfecta para visualizar redes y comunidades, sin embargo, era la única función que le dábamos, los datos que obteníamos después de bajarlos, los importábamos a otra herramienta más especializada en diseño y visualización que se llama **Gephi**,

Podemos aprovechar el recuento para hacer una nota metodológica, **NodeXL**, trabaja sobre una plantilla de Excel, esto quiere decir que se pueden importar y exportar datos, desde y hacia otras herramientas, esta capacidad le permite a **NodeXL** ser una herramienta flexible y abierta para trabajar diferentes tipos de datos.

En un artículo muy difundido en redes sociales de la Dra. Verónica Espinoza, ejemplifica cómo **NodeXL** funciona con otras herramientas. *Con **NodeXL** se puede recopilar, analizar y visualizar redes sociales complejas, con los importadores integrados en **NodeXL** como Twitter, YouTube, Wikipedia y Flickr, pero también puede generar redes con datos externos, solo tiene que descargar los datos de otra fuente y preparar los datos*[1].

[1] https://medium.com/@vespinozag/nodexl-network-analysis-can-be-made-simpler-876ce07d46f2

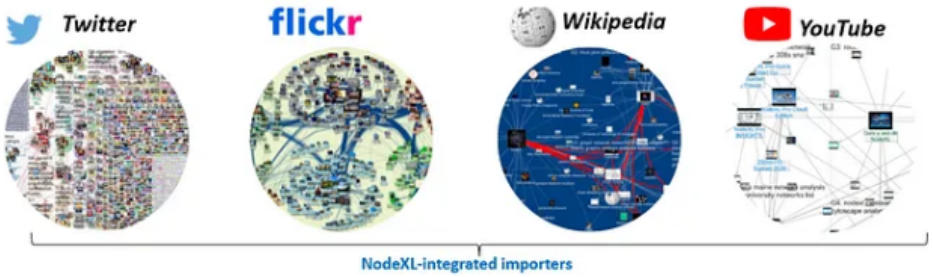


Ilustración 1. Fuente:
https://medium.com/@vespinozag/nodexl-network-analysis-can-be-made-simpler-876ce07d46f2

En la ilustración 1 podemos observar a partir de diferentes redes y visualizaciones creadas por su autora, cómo **NodeXL** tiene flexibilidad para interactuar con datos que provienen de diferentes ambientes y particularmente de redes sociales o medios sociales -*social media* en inglés-, si bien está diseñadas para trabajar con datos de redes sociales, también se pueden trabajar datos que vienen de otro tipo de fuente como transcripciones de entrevistas o grupos focales, para ello hay que trabajar las plantillas de NodeXL, para tales fines se puede consultar un tutorial intitulado en la página de la **SMRF** intitulado *Tutorial Redes Semánticas con NodeXL*[2].

En la página de la Social Media Research Foundation hay tutoriales y recursos muy valiosos para el uso de NodeXL y el aprendizaje de Análisis de Redes Sociales.
Página principal de la SMRF: https://www.smrfoundation.org/
Tutoriales: https://www.smrfoundation.org/nodexl/tutorials/
Asimismo, la página de historias en Medium de la Dra. Verónica Espinoza ofrece diferentes entradas con herramientas que pueden complementar el uso de NodeXL: https://medium.com/@vespinozag

[2] https://www.smrfoundation.org/wp-content/uploads/2022/05/NodeXL-Tutorial-Redes-Semanticas-con-NodeXL-Veronica-Espinoza-Gonzalez-18-febrero-2022.pdf

A finales del 2021, la consultoría **NetHabitus** que fundamos la Dra. Verónica Espinoza y yo, abrió su primer **Curso de Antropología Computacional** dirigido al público en general -principalmente profesionistas- que ya estaban o deseaban iniciarse en el mundo de la investigación digital utilizando herramientas específicas para bajar datos de medios sociales y visualizar y analizar redes sociales. El curso se repitió otras tres veces en vivo en formato remoto y actualmente se encuentra disponible en formato online.

Página principal de NetHabitus: https://www.nethabitus.org/
Página del Curso Antropología Computacional:
https://www.nethabitus.org/cursos/antropolog%C3%ADa-computacional
Página de los diferentes cursos que ofrece NetHabitus:
https://www.nethabitus.org/cursos

Asimismo, en ese mismo espacio de tiempo Marc Smith, director de la **SMRF**, nos invitó a colaborar en la Fundación para llevar a cabo eventos de divulgación y difusión de **NodeXL** dirigidos al público de habla hispana.

Gracias a la colaboración que iniciamos con la **SMRF**, tuvimos la oportunidad de ofrecer licencias de prueba de **NodeXL PRO** funcionales por 30 días para los alumnos participantes en los diferentes cursos, que se han extendido a cursos de **Netnografía** y **Antropología del Consumo.**

Hablar de **NodeXL**, para qué sirve, qué es lo que hace y cuáles son sus atributos y beneficios en el mundo de la investigación actualmente, necesariamente nos remite a otros temas importantes, es decir, necesitamos comprender el contexto en el cual hacemos investigación hoy en día, y ese contexto en gran medida tiene que ver con lo que se ha llamado la “Revolución digital”, simplemente no podemos entender el terreno en el que nos queremos mover en términos de investigación digital si dejamos de lado el momento histórico por el cual pasamos en términos de desarrollo tecnológico.

Asimismo, al ser una herramienta de trabajo diseñada para visualizar y analizar redes sociales, es prácticamente imposible escribir un texto de esta naturaleza y no adentrarse, aunque sea de manera introductoria al tema del Análisis de Redes Sociales (ARS), entender algo de su historia, sus conceptos básicos y cómo se aplica nos ayuda en ampliar nuestra perspectiva sobre cómo podemos aplicar **NodeXL** a nuestros proyectos de investigación ya sean académicos o aplicados al marketing y los negocios.

Establecer las coordenadas de la discusión

Actualmente, el mundo de la investigación cualitativa ha dado un giro -en cantidad y calidad- muy importante hacia los contextos humanos digitales, eso quiere decir que nuestros mundos sociales, así como nuestros hábitos y comportamientos son cada vez más mediados por las tecnologías digitales, y por lo tanto, comprender y analizar estas experiencias requieren nuevas herramientas y recursos tecnológicos que nos permitan dar cuenta de la cultura digital y lo que ésta significa para las personas y la sociedad en el mundo contemporáneo, el cual se caracteriza por el incremento exponencial de la capacidad que tenemos, con los medios tecnológicos digitales, para producir, distribuir y almacenar información. Información que todos, como sociedad interconectada, somos responsables de generar.

Una referencia reciente y destacable en torno a la intersección entre estas dos áreas es el reporte de la UNESCO “New Horizons in Digital Anthropology” (2022), en este documento se da cuenta de cómo la antropología, la cual ha migrado de una praxis eminentemente académica al ámbito comercial y empresarial en muy pocos años, ha encontrado en el espacio digital una expresión que requiere de un balance entre los grandes datos y los datos gruesos (thick data).

En dicho volumen se establece que *La Antropología Digital se enfoca en esta intersección entre la tecnología y los humanos, examinando lo cuantitativo y cualitativo, utilizando Big data y thick data, lo virtual y lo real.*

Tradicionalmente la antropología ha sido una disciplina enfocada en lo que Clifford Geertz ha llamado “descripción densa”, es decir, descripciones a profundidad a partir de la observación directa en espacios y contextos con las personas, en la historia de la disciplina estas observaciones, o etnografías, el cual es el término metodológico correcto, se llevaban a cabo mediante largas estadías en campo de los etnógrafos con las culturas estudiadas, sin embargo, actualmente, las etnografías también se han adaptado al ritmo de las nuevas necesidades de investigación y aunque en la academia se exigen estadías más largas, en el campo de práctica profesional aplicada a marketing y consultoría, la limitante de los tiempos y muchas veces los recursos han llevado a los profesionales a adaptar la metodología a tiempos cortos y con un enfoque muy específico en las necesidades de cada investigación.

Con la llegada de la antropología digital, los antropólogos no hemos dejado de lado la etnografía tradicional, sin embargo, el advenimiento de los grandes datos y la Revolución Digital suponen un cambio de paradigma no solo para la antropología sino para las ciencias sociales en general.

Sin embargo, los mundos digitales, al igual que los mundos sociales y culturales convencionales, son materia de investigación antropológica desde el momento en que son espacios donde las personas crean comunidad, socialización, interacciones, lazos significativos, redes sociales y por lo tanto cultura, y esta tendencia a incorporar lo digital en nuestra vida cotidiana tiene mayor peso en las nuevas generaciones que básicamente es donde construyen su vida social y personal desde edades tempranas.

Es en este contexto donde las ciencias sociales y específicamente la antropología encuentra nuevos desafíos, es decir, aprovechar las grandes cantidades de datos que en muchos sentidos dicen algo sobre quiénes somos en esta era representada por la hiper-conectividad y la hiper-información, y traducir estos datos o esta información, que generalmente se mal entiende como información cuantitativa, y no es necesariamente así, ya que el big data le sirve a los científicos sociales para encontrar patrones y relacionar estos datos en diferentes niveles, como el nivel macro-social con el nivel micro-social, en otras palabras, los grandes datos permiten

visualizar conexiones antes difícil de elaborar (ya que no se contaba con datos y ´poder computacional) y visualizar estructuras complejas en forma de redes sociales, de ahí que el Análisis de Redes Sociales (ARS) es un recurso teórico metodológico que nos permite visualizar, describir, analizar e interpretar estas estructuras sociales antes imposible de ver.

Como lo muestra la ilustración 2, en este artículo proponemos a manera de hipótesis que es mediante el Análisis de Redes Sociales (ARS) que se superan muchas de las contradicciones entre el big data y thick data, dos áreas que podrían clasificarse como cuantitativo/cualitativo pero que bajo esa denominación se pierde el valor y la profundidad de ambas partes.

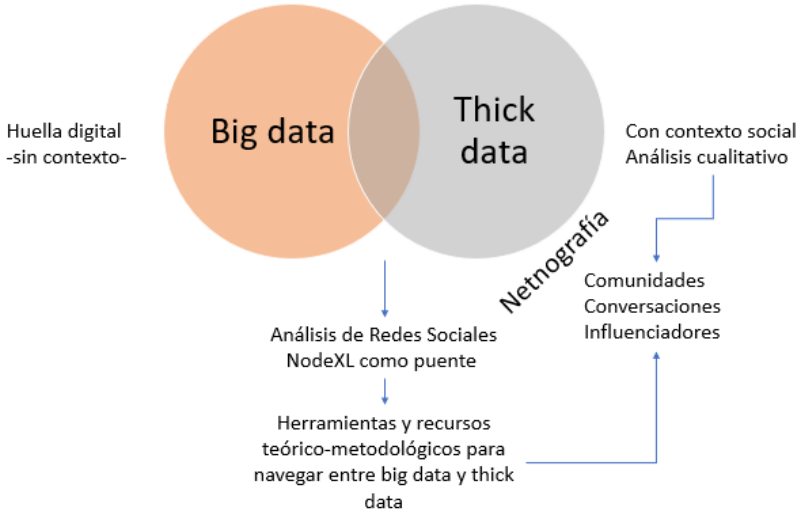


Ilustración 2. Autor: Marcos de Colsa

Así mismo, creemos que el mejor método de investigación que se ha desarrollado para incorporar esta perspectiva es la **Netnografía**, metodología desarrollada por Robert Kozinets desde hace ya varias décadas y que se diferencia de abordajes similares como la etnografía digital o etnografía online en diferentes aspectos, uno de ellos es que hace precisamente énfasis en la huella digital (término que para el análisis cualitativo funciona mejor que “big data”) y particularmente la huella digital en medios sociales o como llamamos en español “redes sociales”, que es diferente al método que reconocemos como Análisis de Redes Sociales (ARS), ya que por

“redes sociales” generalmente se entiende como sinónimo de plataformas como Facebook, Twitter, Instagram, etc.

Y es en esta intersección donde **NodeXL** sirve como recurso metodológico, que es más que una herramienta, es todo un sistema para analizar y visualizar redes (sociales) y hacer puentes entre el mundo de los “grandes datos” y los “datos gruesos” ya que **NodeXL** nos permite detectar comunidades, subcomunidades y actores sociales relevantes o “influenciadores” dentro de una red y los cuales son los agentes que “activan” una comunidad a partir de su actividad, que pueden ser comentarios o publicaciones con diferentes formatos (imagen, video, emojis, texto, audio, etc.). Es a partir de la actividad de los influenciadores que podemos detectar la información cualitativa e indagar más a profundidad para explorar los significados de las conversaciones relevantes, analizar e interpretar dichos significados a manera de “descripción densa”.

La Revolución Digital

Para entender con mayor profundidad y claridad el escenario que planteamos, la intersección entre el big data, el thick data, las redes sociales y el uso de herramientas como **NodeXL**, es importante tomar en cuenta el contexto tecnológico, histórico, social y cultural que vivimos actualmente y que muchos denominan la Cuarta Revolución Industrial o la Revolución Digital.

La Revolución Digital se ha dado en nuestras vidas de forma muy rápida y profunda, ha impactado la vida cotidiana de todos nosotros, prácticamente de la noche a la mañana y sin darnos cuenta, nuestros mundos sociales se volvieron digitales.

En la ilustración 3 podemos observar cómo en menos de una década la tecnología digital móvil se incorporó al mundo social contundentemente, la imagen compara el mismo evento en el mismo lugar sólo ocho años después, y aunque la escena se refiere a un ritual específico de la Iglesia Católica (el anuncio de 2005 del Papa Benedicto XVI y el anuncio de 2013 del Papa Francisco), es un microcosmos y un reflejo del macrocosmos social.

Ilustración 3. Fuente:
https://www.businessinsider.com/vatican-square-2005-and-2013-2013-3?r=MX&IR=T

Lo que vemos en la imagen es realmente el punto focal de la Revolución Digital, mediante nuestros dispositivos digitales dejamos rastros, es decir grandes cantidades de datos e información en torno a las actividades cotidianas que llevamos a cabo, ya sea una fotografía, un mensaje, un video, una nota de voz, todo esto se registra en un ambiente virtual y en conjunto como sociedad creamos grandes cantidades de datos (Big Data).

Estos grandes datos conforman patrones de comportamiento en torno a los diferentes grupos sociales o comunidades de cualquier tipo, es decir, y los patrones que se derivan de estos rastros digitales (o huella digital) dicen algo o más bien mucho de quienes somos como sociedad, qué hacemos, cuáles son nuestras preferencias de consumo, a qué lugares viajamos, con quiénes preferimos socializar en internet o en redes sociales, cómo nos comportamos generacionalmente, qué

música nos gusta y una infinidad más de datos que generamos desde lo individual y que cruza diferentes niveles micro y macro-social.

A partir de estos rastros digitales, los científicos sociales tienen cada vez más recursos para ver y analizar patrones de comportamiento antes imposibles de identificar, por ejemplo, cómo se genera el contagio en medio de una pandemia, cómo fluye la información de diferente tipo, social, económica, política, cultural, etc., y cuál es el comportamiento en torno al flujo de información, cómo responden las comunidades, si hay polarización por ejemplo como ha sido en el caso de las vacunas y los anti-vacunas para el COVID-19, cómo surgen ciertos agentes sociales que fungen como influenciadores en ciertas comunidades y cómo estos influenciadores también juegan un rol determinante en el flujo de información (o desinformación) en diferentes temas y aspectos importantes de la sociedad.

La ilustración 4 muestra cómo actualmente (datos del 2020) hay 5.190 millones de usuarios en dispositivos móviles, lo cual cubre el 67% de la población. Las personas en promedio invierten 3 horas y 40 minutos diarios por usuario, y en ese tiempo el 50% son en aplicaciones de medios sociales y comunicación, 21% en aplicaciones de videos y entretenimiento, 19% en otras aplicaciones y 9% aplicaciones de videojuegos.

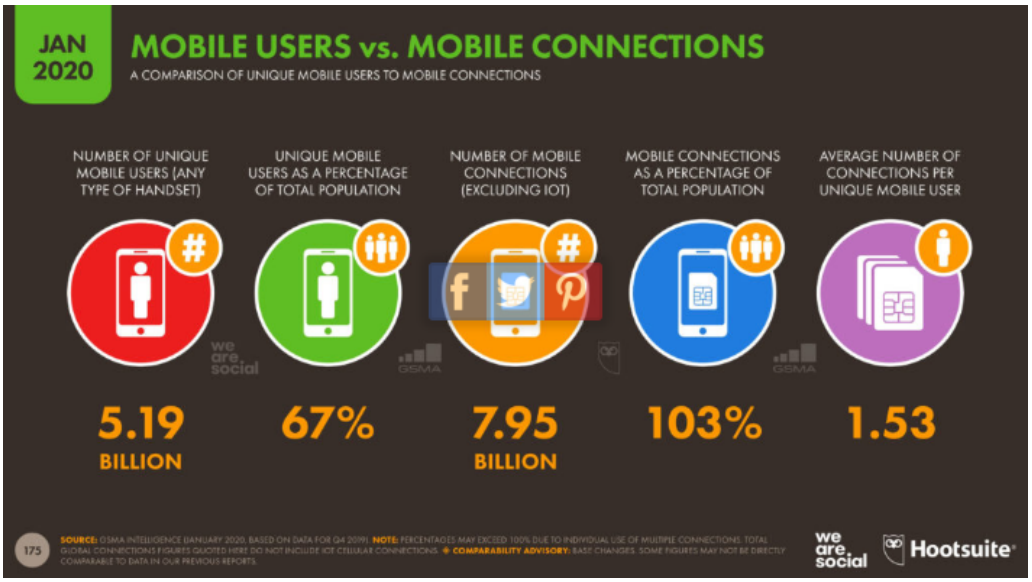


Ilustración 4. Fuente: https://yiminshum.com/mobile-movil-app-2020/

Generalmente para entender mejor este fenómeno, su naturaleza, alcance y evolución, que implica darle sentido a la cantidad de números involucrados y que difícilmente podemos comprender ya que implica un fenómeno que se conoce como “crecimiento exponencial”, quienes han estudiado el desarrollo tecnológico de los últimos 50 años se refieren a la Ley de Moore.

La ley de Moore se ha convertido en el arquetipo de lo que significa la transformación digital, la cual predice el continuo crecimiento exponencial del poder de procesamiento de las computadoras, elaborada por el cofundador de Intel, Gordon Moore en 1965. Esta ley predice una duplicación en el rendimiento de los chips de computadora aproximadamente cada dieciocho meses a medida que los transistores se vuelven más rápidos y pequeños.

Si bien la Ley de Moore se remonta ya a más de medio siglo -y ya se ha hablado del fin de la Ley de Moore- y la tecnología computacional que ahora tenemos se empezó a gestar desde la Segunda Guerra Mundial, sólo con el surgimiento de computadores personales, portátiles y económicamente accesibles y un ecosistema digital como Internet, la Revolución Digital ha tomado forma.

Para entender la Ley de Moore:

- Video: Michio Kaku: modificando la ley de Moore y las computadoras de la era posterior al silicio | Gran pensamiento: https://www.youtube.com/watch?v=bm6ScvNygUU
- Video la ley de Moore no ha muerto (Jim Keller) | Clips de podcasts de IA: https://www.youtube.com/watch?v=c01BlUDIlK4

En la primera revolución industrial, los principales países industrializados utilizaron motores de vapor para mecanizar la producción, sustituyendo a las formas de producción tradicionales.

En la Segunda Revolución Industrial, los avances en materia de energía, materiales, productos químicos y medicina fueron cruciales. Para Rifkin[3] (2009), el motor de combustión interna y la infraestructura de transporte definieron la Segunda Revolución Industrial del siglo XX, la era del petróleo.

La Tercera Revolución Industrial utilizó la electrónica y la computación para automatizar la producción.

La Revolución Digital -para muchos la Cuarta Revolución Industrial- es un proceso de transformación social, económico y cultural que se está produciendo gracias a la adopción masiva de diversas tecnologías relacionadas con el procesamiento de datos: Inteligencia Artificial, Internet de las Cosas, Big Data, realidad aumentada, plataformas colaborativas, la nube o Blockchain entre otras.[4]

Recursos para entender la Revolución Digital y la Inteligencia Artificial (IA):

- Video La Cuarta Revolución Industrial: https://www.youtube.com/watch?v=-OiaE6l8ysg
- Video ¿De qué es capaz la inteligencia artificial?: https://www.youtube.com/watch?v=34Kz-PP_X7c&t=5s
- Video: Cómo cambia nuestra sociedad la inteligencia artificial: https://www.youtube.com/watch?v=CM71Zu7E00E&t=39s

Para los científicos sociales, el surgimiento tanto de un ecosistema digital como Internet y el desarrollo de computadores portátiles cada vez más potentes básicamente representó la oportunidad similar a la que en su momento fue el telescopio para la astronomía y la física o el microscopio para la química y la biología.

[3] Rifkin, J. (2010). La civilización empática. Barcelona, Paidós. ¿Cuarta Revolución Industrial? El reto de la digitalización y sus consecuencias ambientales y antropológicas. Joaquín Fernández Mateo
Revista DIECISIETE. 2021. Nº 4. Madrid. ISSN 2695-4427 Versión digital (pp. 31-46)

[4] Revista DIECISIETE. 2021. Nº4. Madrid. ISSN 2695-4427 Versión digital (pp. 11-12)

En un artículo fundacional intitulado *Computational Social Science*[5] los autores mencionan que *Internet ofrece una forma completamente diferente para entender lo que dice la gente, y como se conectan…una ciencia social computacional emergente apalanca la capacidad de recolectar y analizar datos con un nivel sin precedentes amplitud, profundidad y escala*.

Asimismo, en otro artículo, Duncan Watts nota que, *Aunque la economía, sociología, ciencias políticas y antropología han producido una plétora de hallazgos en cuanto al comportamiento social humano, han tenido mucho menos éxito que las ciencias físicas y de la vida en la producción de un marco teórico coherente que puede dar cuenta de sus descubrimientos. Esto no es porque los científicos sociales son menos listos que sus pares en otros campos, sino porque los fenómenos sociales se encuentran entre los más difíciles problemas científicos a resolver*[6].

Y continúa *La llamativa proliferación en la pasada década de comunicación basada en Internet y la interactividad, sin embargo, está empezando a levantar estas restricciones. Por primera vez, nosotros podemos comenzar a observar las interacciones en tiempo real de millones de personas en una resolución que es sensible a los efectos a nivel del individuo. Mientras tanto, cada vez más rápido las computadoras nos permiten simular grandes redes de interacciones sociales*[7].

De acuerdo con Duncan Watts, el resultado de lo anterior es que ha habido un tremendo interés en las redes sociales: miles de papeles y un creciente número de libros han sido publicados en menos de una década, lo que lleva a algunos a anunciar la llegada de una 'ciencia de las redes'.

---

[5] Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A. L., Brewer, D., ... & Van Alstyne, M. (2009). Computational social science. *Science*, *323*(5915), 721-723.

[6] Watts, D. J. (2007). A twenty-first century science. *Nature*, *445*(7127), 489-489.

[7] Watts, D. J. (2007). A twenty-first century science. *Nature*, *445*(7127), 489-489.

Recursos sobre Ciencias Sociales Computacionales:

- Serie de videos en YouTube: UCCSS: University of California Computational Social Science: https://www.youtube.com/playlist?list=PLtjBSCvWCU3rJ_XAqaOr7NoQYWY005WCI (con subtítulos)
- Summer Institute in Computational Social Science: https://www.youtube.com/@summerinstituteincomputati9729 (sin subtítulos)

Big Data Vs. Thick Data

> *[Big data] No es un término muy afortunado, pero para todo lo que concierne a las ciencias sociales, comprender la sociedad y cómo se usa en la sociedad, este fenómeno, también podría reemplazar este término con* ***huella digital*** *que explica mucho mejor de qué se trata*. -Martin Hilbert-

Actualmente hay una visión sumamente polarizada en torno a cómo hacer investigación en los medios digitales, la narrativa es que se hay una tradición que viene de los métodos cuantitativos y otra que viene de los métodos cualitativos, los grandes datos o big data tiene una connotación inevitablemente asociada al mundo de lo cuantitativo, sin embargo la lógica del big data no es llevar a cabo estadísticas o sacar medias o tendencias, sino encontrar patrones en el comportamiento humano a partir de los rastros digitales, es decir, una lógica que comparte con el mundo cualitativo más que con el terreno de los números.

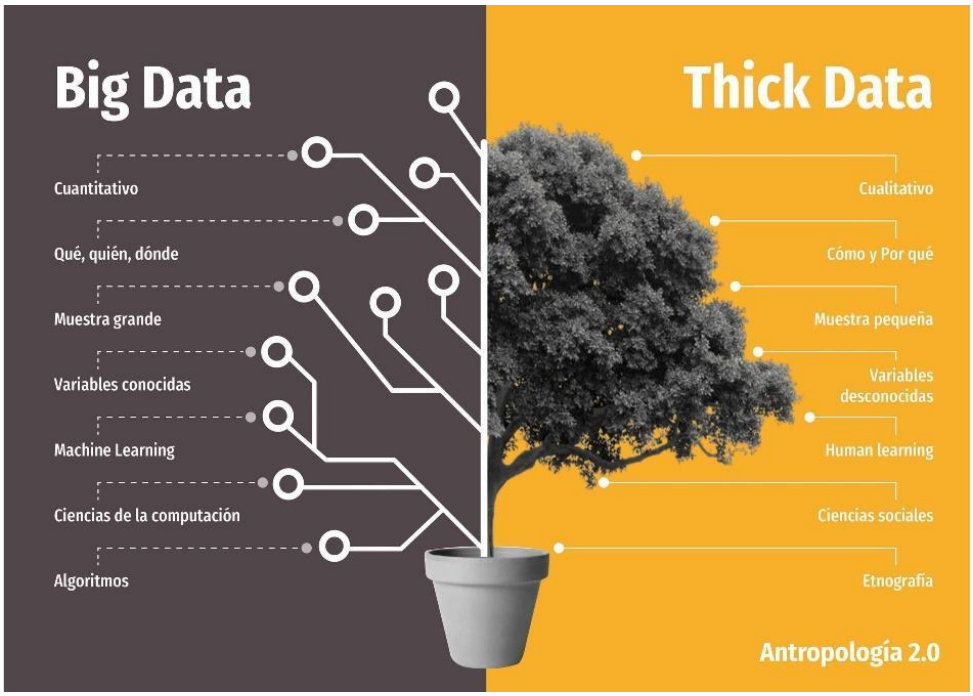


Ilustración 5. Fuente: https://blog.antropologia2-0.com/es/que-es-thick-data/

Sin embargo, el big data sí tiene una connotación de cantidad, aunque el grueso de esos datos -que en muchos casos son metadatos- no son accesibles para la mayoría de los investigadores que trabajan desde su computadora portátil, ya que son datos que pertenecen a las grandes empresas tecnológicas como Google o de comunicación y telefonía celular que proveen los servicios, acceder a ese universo de datos se requieren equipos de computación con mayor poder de cómputo a los comerciales y un gran presupuesto y generalmente un equipo multidisciplinario para trabajar y procesar los datos, además de contar con los permisos necesarios de las empresas tecnológicas y/o de comunicación.

Cuando hacemos algo en una computadora, o en un teléfono celular digitalmente, esta información se transfiere a ceros y unos luego se transmiten a través de un cable de fibra óptica o a través del aire a través de ondas de radio o se almacena en un disco duro, pero estos seriales se almacenan allí, se almacenan y dejan una huella atrás, a esta información se le ha llamado “big data” o “huella digital”.

Para muchos investigadores, “big data” es un término pobre. Se ha utilizado en las ciencias para referirse a conjuntos de datos lo suficientemente grandes como para requerir supercomputadoras, aunque ahora se pueden analizar vastos conjuntos de datos en computadoras de escritorio con software estándar.

Las cantidades de datos ahora disponibles son realmente grandes, pero esa no es la característica más relevante de este nuevo ecosistema de datos. “el Big Data es notable no por su tamaño, pero por su relacionalidad con otros datos.[8]”

Su valor proviene de los patrones que se pueden derivar al hacer conexiones entre piezas de datos, sobre un individuo, sobre individuos en relación con otros, sobre grupos de personas, o simplemente sobre la estructura de la información en sí, en otras palabras, el Big Data está fundamentalmente interconectado (networked).

En un muy difundido esquema de la IBM -ilustración 6-, los científicos de datos de esta compañía establecieron que el big data se debe entender y considerar a partir de lo que llamaron las 4 V´s del big data.

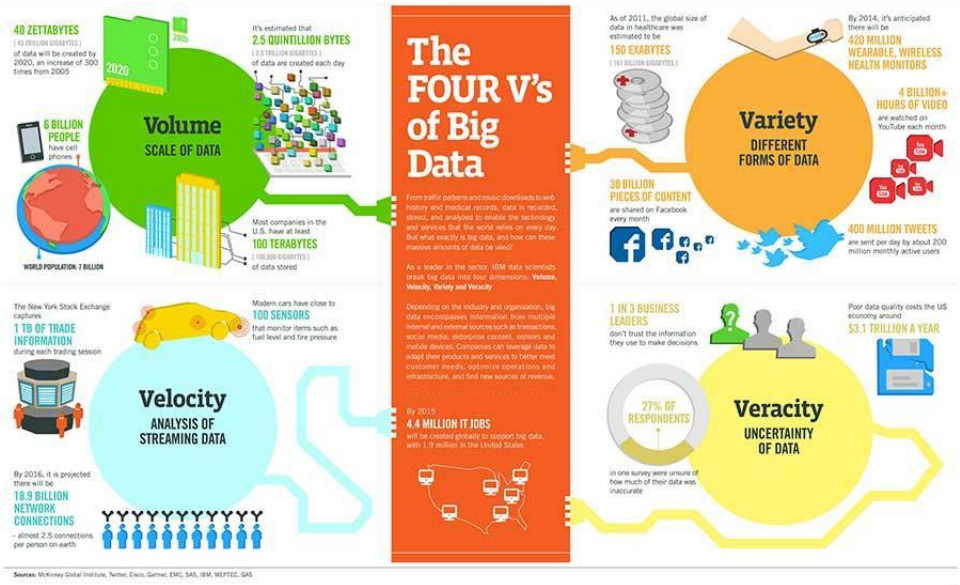


Ilustración 6, las 4 V´s del Big Data de IBM. Fuente:
https://opensistemas.com/en/the-four-vs-of-big-data/

[8] Boyd, D., & Crawford, K. (2011, September). Six provocations for big data. In *A decade in internet time: Symposium on the dynamics of the internet and society*

Se puede decir entonces que el entorno Big Data debe tener estas cuatro características básicas: Volumen, Variedad, Velocidad y Veracidad.

- **Volumen**: Criterio de cantidad (grandes, pero inevitablemente con sesgo)
- **Variedad (data-fusión):** Diversidad de tipos de datos (sitios, fuentes, formato, etc.)
- **Velocidad:** Para generarse y analizarse. Es dinámica. Se produce en tiempo real.
- **Veracidad (huella digital):** Verificables, registrados de forma automática, muchas veces de forma no intencional

Además de las 4 V's de IBM es importante tomar en cuenta que también hay dos grandes tipos de huella digital: pasiva y activa. Las pistas pasivas se dejan a través de la interacción con una infraestructura, como una red de telefonía móvil, que produce entradas en registros de ubicación y generalmente es el tipo de información al cuál los investigadores no tenemos acceso.

Por su lado, la huella digital activa, que son las impresiones activas que provienen de los propios usuarios cuando exponen datos en fotos, mensajes, etc., es el tipo de datos que trabajamos cuando hacemos investigación en medios digitales como Facebook, Instagram o Twitter u otras redes sociales.

Con el advenimiento de los grandes datos, los antropólogos comienzan a cuestionar qué se está perdiendo a través de la obsesión global con la ciencia de datos: ¿conocimiento cultural, intuición humana, empatía?[9] . Los datos gruesos son una respuesta antropológica a los grandes datos, que incorporan ricos matices humanos a escala y defienden la necesidad de su integración en las prácticas de análisis de datos.

[9] *NEW HORIZONSIN DIGITAL ANTHROPOLOGY, UNESCO, 2022*

En el 2013, Tricia Wang escribe un artículo muy influyente intitulado *Big Data necesita datos gruesos*, en donde establece que *El análisis de datos gruesos se basa principalmente en el poder del cerebro humano para procesar una "N" pequeña, mientras que el análisis de datos grandes requiere poder computacional (por supuesto, con humanos escribiendo los algoritmos) para procesar una "N" grande. Big Data revela información con un rango particular de puntos de datos, mientras que Thick Data revela el contexto social y las conexiones entre los puntos de datos. Big Data entrega números; Los datos gruesos ofrecen historias. Big data se basa en el aprendizaje automático; Los datos gruesos se basan en el aprendizaje humano*[10].

Por lo tanto, el concepto de datos gruesos es un sinónimo de datos observacionales recopilados y analizados etnográficamente en la tradición de Clifford Geertz (1977), quien refirió cómo las descripciones gruesas del comportamiento humano incluyen la recopilación detallada de datos y el análisis del contexto en el que ocurre un comportamiento.

De acuerdo con una definición reciente, los datos gruesos se definen por su complejidad contextual que permite al investigador reflexionar sobre cómo y por qué las personas hacen lo que hacen. *Se puede definir "thick data" como información cualitativa compleja sobre un grupo de sujetos, tradicionalmente obtenida a través de la etnografía, pero ahora cada vez más a través de software especialmente diseñado, revelando contextos, emociones, intimidades.* (UNESCO, 2022).

Este tipo de datos a menudo se recopilan mediante un software especialmente diseñado que recopila datos cualitativos de conocimiento humano o recopila datos de fuentes abiertas (como las redes sociales), es justo en este terreno donde **NodeXL** nos sirve como puente entre el mundo de los grandes datos y el thick data mediante la visualización y el Análisis de Redes Sociales.

Aunque el artículo de Tricia Wang, tal vez sin habérselo propuesto ha contribuido a la polarización del discurso entre big data y thick data, sin embargo, ella misma propone oportunidades de interacción entre las dos áreas y una de ellas es el Análisis de Redes Sociales.

[10] http://ethnographymatters.net/blog/2013/05/13/big-data-needs-thick-data/

Recursos para entender el Big Data:

- Video: Google, Facebook, Amazon - El poder ilimitado de los consorcios digitales | DW Documental: https://www.youtube.com/watch?v=A3cGMNxRNJ0
- Video: Amazon, Jeff Bezos y la colección de datos | DW Documental: https://www.youtube.com/watch?v=UzGemfwaTT8
- Video: Documental "Big Data: el valor de nuestra información" de Modesto Sierra UNED: https://www.youtube.com/watch?v=xDIV1jCW7n0

Análisis de Redes Sociales

¿Qué es una red social? Veamos algunas definiciones:

En un sentido muy básico, una red social es un conjunto organizado de personas que consta de dos tipos de elementos: los seres humanos y las conexiones entre ellos. (Christakis, N. A., & Fowler, J. H. 2009)[11].

La estructura elemental de cualquier red está constituida por elementos en relación. Para las redes sociales, los elementos son normalmente personas y las relaciones pueden verse como transacciones o intercambios. (Wolfe, A. W. 2011)[12].

Un grupo de personas que están conectados por relaciones sociales particulares, como el parentesco, la amistad, el trabajo, un pasatiempo compartido o un interés común, o intercambiar cualquier tipo de información puede considerarse una red social. (Kozinets, R. V. 2010)[13].

Una red social consiste en un conjunto de actores, que pueden ser entidades arbitrarias como personas u organizaciones, y un conjunto de relaciones entre ellos. (Brandes, U., & Wagner, D. 2000)[14].

Las diferentes definiciones aquí expuestas provienen de diferentes tradiciones, por ejemplo, la primera más relacionada a la sociología, la segunda con la antropología social, la tercera con el marketing y la última de la ciencia computacional, sin embargo, todas tienen algo en común, hacen referencia a dos elementos básicos: 1) actores sociales o personas y 2) relaciones entre ellos.

---

[11] Christakis, N. A., & Fowler, J. H. (2009). Connected: The surprising power of our social networks and how they shape our lives. Little, Brown Spark

[12] Wolfe, A. W. (2011). Anthropologist view of social network analysis and data mining. Social Network Analysis and Mining, 1(1), 3-19.

[13] Kozinets, R. V. (2010). Netnography: Doing ethnographic research online. Sage publications

[14] Brandes, U., & Wagner, D. (2000). Contextual visualization of actor status in social networks. In “Data Visualization 2000” (pp. 13-22). Springer, Vienna.

Desde una perspectiva más amplia el Análisis de Redes Sociales (ARS) es un método analítico que se centra en las estructuras y patrones de relaciones entre los actores sociales en una red. En el análisis de redes sociales, hay dos unidades principales de análisis.

1. Los actores sociales que nos interesan se llaman 'nodos' y
2. La relación entre ellos se llama ‘ejes o vínculos'.

De hecho, el Análisis de Redes Sociales es ahora un cuerpo teórico metodológico que se ha conformado históricamente a partir de diferentes tradiciones científicas.

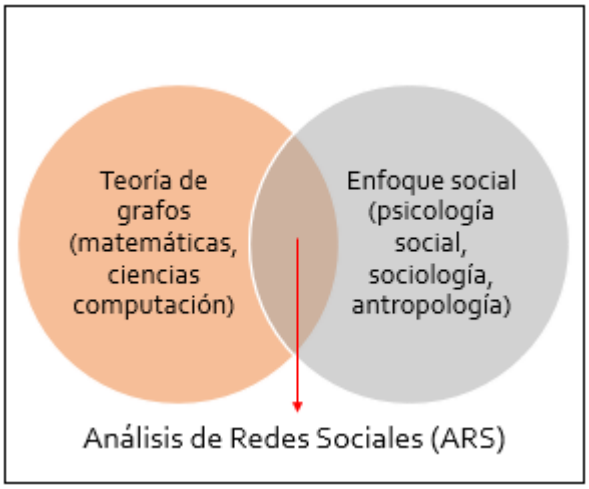


Ilustración 7. Autor: Marcos de Colsa

El ARS tiene sus fundamentos en la sociología, la sociometría y la teoría de grafos, y en la línea estructural-funcionalista de los "antropólogos de Manchester, que se basaron en ambas vertientes para investigar la estructura de las relaciones "comunitarias" en las sociedades tribales y de aldea".

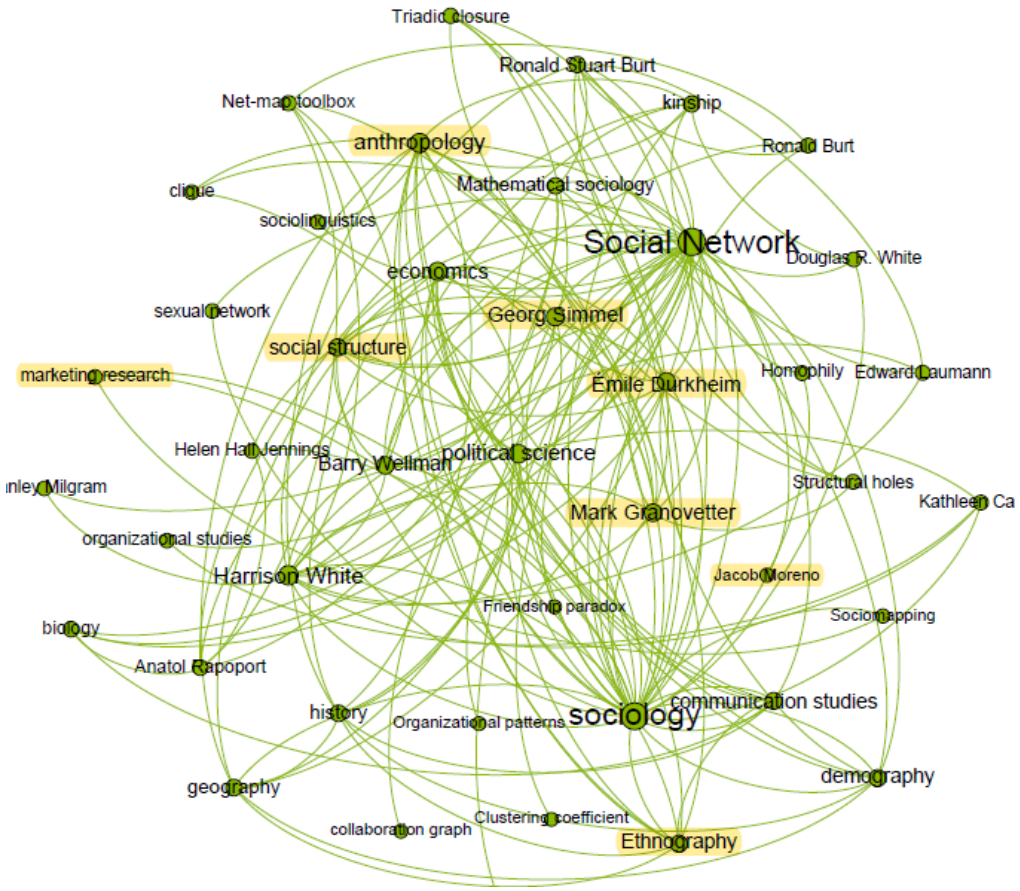


Ilustración 8. Autor: Marcos de Colsa. Grafo elaborado a partir de datos de Wikipedia usando **NodeXL** y Gephi.

En la ilustración 8, podemos observar cómo el concepto de red social se relaciona con algunos campos y autores que nos hablan tanto de su historia como de su estado actual, uno de los nodos (actores) más relevantes en la gráfica es “Sociology”, dando cuenta del rol determinante de algunos sociólogos en la conformación de este paradigma.

Emile Durkheim (1893-1964) describió las sociedades tradicionales en el que la solidaridad mecánica vinculó a individuos similares con regulaciones represivas. Esto lo distinguió de las sociedades modernas, en el que una división del trabajo llevó a los individuos a formar cooperativas basadas en la solidaridad orgánica.

Georg Simmel (1858-1918) fue un filósofo, sociólogo y crítico alemán. Simmel expresó la creencia central que subyace en el análisis moderno de las redes

sociales, para Simmel, la sociología era el estudio de los patrones de interacción social.

Desde el campo de la psicología Jacob Levy Moreno (1889-1974, es una de las figuras clave en el desarrollo de una ciencia de redes. Levy Moreno partió del psicodrama para desarrollar lo que se conocería como "sociometría". *El psicodrama pone al paciente sobre un escenario, donde puede resolver sus problemas con la ayuda de unos pocos actores terapéuticos. Es tanto un método de diagnóstico como de tratamiento.* (Moreno, 1946).

Las mediciones sociométricas comenzaron con cosas como éstas: ¿cuánto "tiempo" pasa el actor A con el actor B? [...] O ¿cuál es la "distancia espacial", cerca o lejos, en pulgadas, pies o metros, entre los actores A, B, C y D en la misma situación?, ¿y qué efectos tiene la proximidad o la distancia sobre la conducta y la actuación? O ¿con cuánta frecuencia aparecen dos actores simultáneamente en la escena y con cuánta frecuencia se van juntos? (Moreno 1953).

El conjunto de esos canales constituía una especie de diagrama del flujo de la energía afectiva en un grupo social. En su contexto de origen el sociograma era inseparable del psicodrama. Levy Moreno pensaba que el potencial terapéutico de sus representaciones escénicas derivaba de las interacciones entre los actores.

De acuerdo con el antropólogo Carlos Reynoso (2011)[15], Levy Moreno introdujo cinco ideas claves en la construcción de imágenes de redes sociales:

1. dibujó grafos;
2. dibujó grafos dirigidos;
3. utilizó colores para trazar multigrafos;
4. varió las formas de los puntos para comunicar características de los actores sociales;
5. y, mostró que las variaciones en la ubicación de los puntos podían usarse para subrayar importantes rasgos estructurales de los datos

[15] Reynoso, C. (2011). *Redes sociales y complejidad: Modelos interdisciplinarios en la gestión sostenible de la sociedad y la cultura*. Buenos Aires: Sb.

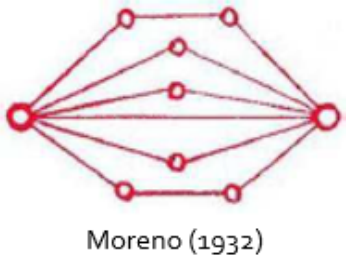
	Uno de los primeros grafos dibujados por Levy Moreno. Fuente: Reynoso (2011)

Ilustración 9.

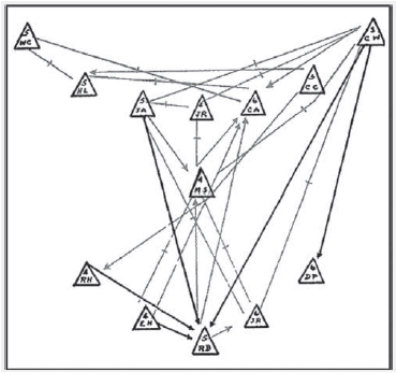
	El primer diagrama de red social de Jacob Moreno de relaciones positivas y negativas entre los miembros de un equipo de fútbol. Publicado originalmente en Moreno, J.L., 1934. ¿Quién sobrevivirá?

Ilustración 10.

Moreno desarrolló la sociometría aplicando estos diagramas en estudios de relaciones entre miembros de un equipo de fútbol. Estos diagramas revelaron patrones de amistad y animosidad.

Desde la antropología social, un momento crucial para en la historia del desarrollo del Análisis de Redes Sociales fue la llamada Escuela de Manchester que tuvo su auge entre 1955-1975 y fue liderada por el antropólogo Max Gluckman (1911-1975).

La Escuela de Manchester se caracteriza por:

- Ser una alternativa opuesta a los planteamientos sincrónicos y estáticos de la antropología sociocultural inglesa, de tono estructural-funcionalista.

- Se hicieron los primeros aportes británicos a la antropología urbana.
- Se propusieron teorías de la dinámica y el cambio y se usaron por primera vez redes antropológicas.
- Se articularon por primera vez los “estudios de casos”.
- Llamada a veces antropología interaccional, enfatizó las redes centradas en Ego (o redes personales) constituidas a nivel urbano.

Dentro de esta Escuela, según Carlos Reynoso, el creador del concepto de red social con ese nombre exacto fue el antropólogo John Barnes (1954), quien promovió el pasaje de una concepción metafórica a una afirmación conceptual sobre relaciones sociales.

Otro personaje importante, que también vemos como nodo en la ilustración 8, es Mark Granovetter (1943) quien es un sociólogo estadounidense de la Universidad de Stanford. En 1973, publicó un artículo influyente intitulado *La fuerza de los lazos débiles*, los lazos fuertes pueden unir a los individuos en grupos, pero los lazos débiles unen a los grupos en la sociedad en general y son cruciales para la propagación de la información.

Granovetter argumentó que estas conexiones más débiles actúan frecuentemente como puentes de un grupo a otro y, por lo tanto, juegan un rol crítico.

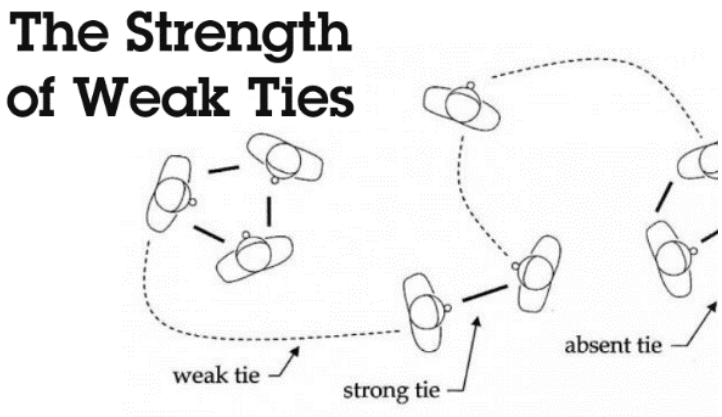
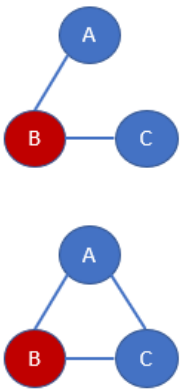


Ilustración 11. Fuente: https://visiblenetworklabs.com/2019/08/21/the-strength-of-weak-ties-paper/

El autor encontró que de las 54 personas entrevistadas que habían encontrado trabajo a través de contactos sociales, 16,7% lo logró a través de lazos fuertes, 55,7% mediante lazos de fuerza intermedia y 27,6% merced a un vínculo débil. Eran amigos lejanos o amigos de amigos que pasaron sus nombres a un empleador o que pasaron información sobre trabajos al posible empleado.

El fenómeno que descubrió Granovetter se puede explicar en función de cómo se teje una red de confianza mediante un "cierre triádico", como se puede ver en la ilustración 12.



- En una red de confianza, es probable que se desarrolle un "cierre triádico" debido a la propiedad transitiva.
- Si un nodo A confía en el nodo B, y el nodo B confía en el nodo C, el nodo A tendrá la base para confiar en el nodo C.
- En una red social, se produce un cierre triádico fuerte porque hay una mayor oportunidad para que los nodos A y C tengan como vecino común a B y por lo tanto crear al menos lazos débiles.
- El nodo B también tiene el incentivo de unir A y C para disminuir el estrés latente en dos relaciones separadas.

Ilustración 12. Autor: Marcos de Colsa

En un estudio reciente, publicado en la revista Science, intitulado *A causal test of the strength of weak ties*[16], los autores llevaron a cabo un experimento donde

[16] Rajkumar, K., Saint-Jacques, G., Bojinov, I., Brynjolfsson, E., & Aral, S. (2022). A causal test of the strength of weak ties. *Science*, *377*(6612), 1304-1310.

básicamente hacen una actualización de la teoría de Granovetter, realizado con el algoritmo “Personas que quizás conozcas” de LinkedIn, que recomienda conexiones a los usuarios, los experimentos mostraron que los lazos débiles aumentan las transmisiones de trabajo, pero solo hasta cierto punto. Los autores muestran que los lazos más débiles tuvieron el mayor impacto en la movilidad laboral, mientras que los lazos más fuertes tuvieron el menor. *Juntos, estos resultados ayudan a resolver la aparente "paradoja de los lazos débiles" y proporcionan evidencia de la teoría de la fuerza de los lazos débiles*[17]*.*

En síntesis, es un campo de estudio que sigue la tradición de las investigaciones sociales que se han centrado consistentemente en las relaciones sociales que vinculan a los individuos más que en los propios individuos, este tipo de investigación que examina los vínculos entre los objetos de estudio se denomina *estructural,* no se limita al estudio de las relaciones sociales humanas y se encuentra presente en casi todos los campos de la ciencia.

El enfoque de la *red social* se basa en la noción de que el patrón de los lazos sociales en los que los actores están incrustados tiene consecuencias importantes para esos actores, los analistas, entonces, buscan descubrir varios tipos de patrones y tratan de determinar las condiciones bajo las cuales esos patrones surgen y buscan descubrir sus consecuencias.

De acuerdo con Linton C. Freeman en *The development of social network análisis,* (2004) las diferentes tradiciones de las cuales ha derivado el ARS coinciden en el enfoque contemporáneo donde:

1. El análisis de redes sociales está motivado por una intuición estructural basado en lazos que vinculan a los actores sociales,
2. Se basa en datos empíricos sistemáticos,
3. Se basa en gran medida en imágenes gráficas, y
4. Se basa en el uso de medidas matemáticas y/o computacionales modelos.

[17] https://www.science.org/doi/abs/10.1126/science.abl4476

Para entender más cómo funcionan las redes sociales:

- Video Marc Smith | Network Mapping the Ecosystem: https://www.youtube.com/watch?v=kDiGl-2m868&t=6s (con subtítulos)
- Video 2 CCSSCS: Análisis de Redes Sociales (parte 1): https://www.youtube.com/watch?v=oy8YxTshZhI&t=2586s
- Diseño de redes en diferentes ámbitos de las ciencias y humanidades: http://www.visualcomplexity.com/vc/

Netnografía

La Netnografía es una metodología desarrollada por Robert Kozinets en 1995 estudiando las comunidades virtuales de Star Trek y desde entonces ha publicado tres libros y una infinidad de artículos sobre la netnografía y su aplicación en diferentes ámbitos de las ciencias sociales, particularmente en el terreno del marketing y el consumo.

En palabras de Kozinets, *La netnografía se puede definir como un relato escrito resultante del trabajo de campo que estudia las culturas y comunidades que surgen de las comunicaciones en línea, mediadas por computadora y/o basadas en Internet, donde tanto el trabajo de campo como el relato textual están informados metodológicamente por las tradiciones y técnicas de la Antropología cultural.*[18]"

Por lo tanto, netnografía es el nombre dado a un conjunto específico de recopilación de datos relacionados, análisis, prácticas de investigación éticas y representacionales, donde la cantidad de datos recopilados e investigación de

[18] Kozinets, R. V. (1998). On netnography: Initial reflections on consumer research investigations of cyberculture. ACR North American Advances.

observación participante realizadas tienen su origen y se manifiesta a través de los datos compartidos libremente en el Internet, incluidas las aplicaciones móviles.

Este énfasis en cantidades significativas de datos de Internet diferencia la netnografía de enfoques como la etnografía o antropología digitales que son más generales en orientación y, por ejemplo, puede incluir etnografías más tradicionales que examinan cómo las culturas nacionales influyen en el uso de Internet móvil, el uso de tecnología dentro del hogar o las culturas ocupacionales de los desarrolladores de software.

En netnografía***,*** *los investigadores ven los datos cualitativos en línea como indicativos de culturas o comunidades,* utilizando técnicas prevalentes en antropología y sociología, *los investigadores de marketing y consumo pueden estudiar las redes sociales y las comunidades en línea como fenómenos culturales*[19]*.*

El nombre Netnografía, un acrónimo que combina red, Internet y etnografía, siempre ha representado un enfoque cultural en la comprensión de los datos derivados de los datos de las redes sociales, características que el enfoque comparte con todos los demás tipos de etnografía en línea.

Es una metodología que se centra en el estudio de los rastros en línea, algunas netnografías recopilan y analizan exclusivamente rastros en línea, otras se extienden más allá de los rastros en línea a otras formas de recopilación y creación de datos. Sin embargo, el enfoque en los rastros en línea que, cuando se recopilan, se convierten en datos en línea, es un elemento distintivo clave de la netnografía.

Cuando las personas publican imágenes, videos o mensajes de texto en línea, o cuando comentan, comparten o hacen cualquier otra cosa que sea accesible en línea para otras personas anónimas o conectadas en red, lo que dejan son rastros en línea (huella digital). Los rastros en línea pueden ser textuales, gráficos, fotográficos, audiovisuales, musicales y muchas otras cosas más.

---

[19] Kozinets, R. V. (2019). *Netnography: The Essential Guide to Qualitative Social Media Research.* Sage publications

Desde que Kozinets empezó a desarrollar la netnografía, la recopilación e interpretación de estos rastros digitales ha constituido el núcleo de una investigación netnográfica.

En síntesis, se puede decir que la netnografía se caracteriza principalmente por:

- Ser un tipo específico de investigación cualitativa en redes sociales.
- Adapta los métodos de etnografía y otras prácticas de investigación cualitativa a las experiencias culturales que abarcan y se reflejan dentro de las huellas, y sistemas de las redes sociales.
- La netnografía difiere del campo más general de la investigación digital por su énfasis en los rastros digitales, interacciones y socialidades en línea.
- Finalmente, la netnografía es un enfoque pragmático, práctico y a nivel de silla de trabajo para estudiar las redes sociales utilizando una lente cultural.

Al ser una metodología que adapta las técnicas de investigación de la antropología social en medios digitales, con énfasis en el estudio de las comunidades online, hay por lo tanto una relación directa con el Análisis de Redes Sociales y con las posibilidades que **NodeXL** ofrece para conectar el enfoque metodológico de la netnografía y el ARS.

En su libro *Netnography: Redefined*. (2025), Kozinets dedica una sección del libro en describir la relación que se da entre el Análisis de Redes Sociales y la netnografía.

Los analistas de redes sociales buscan describir las redes de relaciones lo más completamente posible, desentrañar los patrones prominentes en tales redes, rastrear el flujo de información (y otros recursos) a través de ellas y descubrir qué efectos tienen estas relaciones y redes en las personas y las organizaciones (o comunidades). Hay, por lo tanto, una superposición considerable con ciertos tipos de netnografía, que pueden centrarse en las relaciones y los patrones estructurados de intercambios de cosas como el lenguaje, los símbolos, el discurso, los valores, el poder y otros recursos simbólicos y materiales

Los netnógrafos también consideran esos recursos, viéndolos en y desde diversos contextos y superpuestos, que podrían incluir como fuentes múltiples y compartidas de significado y también como portadores de conexión interpersonal.

De acuerdo con Kozinets, hay muchas oportunidades de sinergias entre el análisis estructural de las redes sociales y los análisis más identitarios, narrativos, centrados en el curso y el significado de la netnografía.

Hay muchas maneras de conceptualizar el universo de formas de redes sociales para obtener una visión básica de los diferentes tipos de conexión que las personas tienen entre sí en línea. Hay dos formas esenciales e interrelacionadas en que las personas se conectan entre sí: socialmente y a través de temas.

Kozinets cita en su libro el artículo que la **SMRF** publicó junto con el Pew Research Center, *Mapping Twitter Topic Networks: From Polarized Crowds to Community Clusters*, en donde mediante el uso de **NodeXL**, los autores clasificaron, después de analizar una gran cantidad de redes sociales que contienen conversaciones y datos de Twitter, las diferentes estructuras más comunes que se dan al visualizar dichas conversaciones.

Las conversaciones en Twitter crean redes con contornos identificables a medida que las personas responden y se mencionan entre sí en sus tweets. Estas estructuras conversacionales difieren, dependiendo del tema y de las personas que impulsan la conversación. Seis estructuras se observan regularmente: divididas, unificadas, fragmentadas, agrupadas y estructuras radiales hacia adentro y hacia afuera. Estos se crean a medida que las personas eligen a quién responder o mencionar en sus mensajes de Twitter y las estructuras cuentan una historia sobre la naturaleza de la conversación[20]*.*

Estas seis estructuras distintivas, ver ilustración 12, nos informan sobre algunas de las diversas formas que puede tomar la sociabilidad en línea, dependiendo del tema de la conversación, el tipo de conexiones entre actores individuales en la red, las fuentes de información y otros recursos que se utilizan, redes corporativas, transaccionales y sociales que están involucradas, los líderes de la conversación y la estructura de la conversación.

[20] Smith, M. A., Rainie, L., Shneiderman, B., & Himelboim, I. (2014). Mapping Twitter topic networks: From polarized crowds to community clusters.

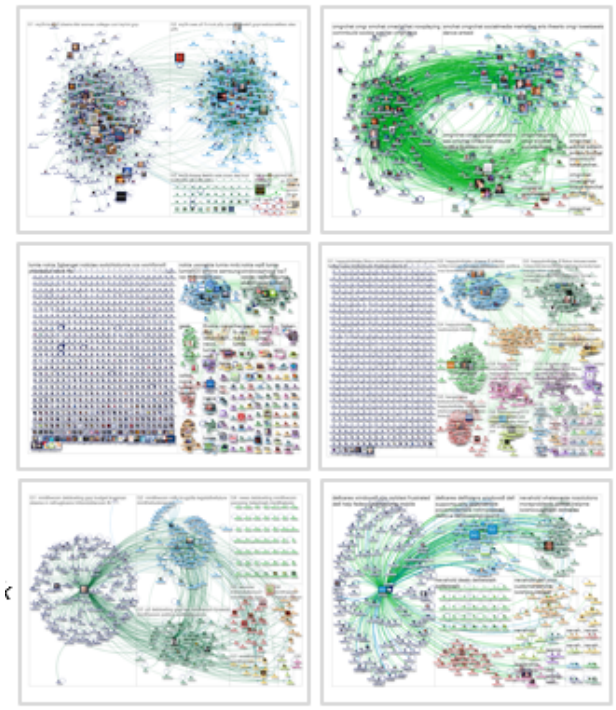


Ilustración 13. Fuente: https://www.pewresearch.org/internet/2014/02/20/mapping-twitter-topic-networks-from-polarized-crowds-to-community-clusters/

Las seis estructuras o “arquetipos” como las llaman los autores son:

1. ***Multitud polarizada:*** Dos grandes grupos que tienen poca conexión entre ellos
2. ***Multitud cohesionada:*** Se caracterizan por personas altamente interconectadas con pocos participantes aislados.
3. ***Clústeres de marcas:*** Cuando se discuten productos o servicios conocidos a menudo hay comentarios de muchos participantes desconectados
4. ***Clústeres comunitarios:*** Algunas fuentes de información y temas encienden múltiples conversaciones, cada una cultivando su propia audiencia y comunidad.
5. ***Red de diffusion:*** Mucha gente repite lo que tuitean organizaciones importantes de medios y noticias

6. ***Red de soporte:*** Las quejas de los clientes para una empresa importante a menudo se manejan mediante una cuenta de servicio de Twitter que intenta resolver y gestionar los problemas de los clientes.

El ARS es un aliado poderoso de la netnografía cundo se busca:

- Aprender sobre la estructura de las comunicaciones de una comunidad.
- Discutir patrones de relaciones sociales o "vínculos“.
- Describir diferentes tipos de relaciones sociales e intercambios entre miembros de una red social identificada.
- Investigar los patrones de comunicación e interacción en línea
- Estudiar los flujos de comunicación y conexión a través de diferentes sitios y redes en línea
- Comparar las estructuras comunitarias y los flujos de comunicación entre diversos tipos de redes sociales, interacciones y experiencias.

Asimismo, de acuerdo con Robert Kozinets, el ARS permite responder preguntas de investigación como:

- ¿Cuál es la estructura de las comunicaciones en un sitio o red en línea en particular?
- ¿Quién se comunica con quién? ¿Quién se comunica más?
- ¿Quiénes son los comunicadores más influyentes en esta red social online?
- ¿Hay un grupo central y un grupo periférico en esta red en particular?
- ¿Cómo se conecta este individuo con otros individuos en y a través de redes sociales?
- ¿Cuáles son los diversos subgrupos en esta red social?
- ¿Cómo fluye la información a través de esta red en particular?
- ¿En qué se diferencia la comunicación en un sitio o plataforma en línea de las comunicaciones en otro sitio o plataforma en línea en términos de quién la usa y qué se comunica?
- ¿Cuáles son los patrones generales en la difusión de información entre una o más redes sociales en línea en particular?

Usando NodeXL para crear puentes

Conceptos básicos para el uso de NodeXL

Vértices

Los vértices, también llamados nodos, agentes, entidades o elementos, pueden representar muchos tipos de cosas. A menudo representan a personas o estructuras sociales como grupos de trabajo, equipos, organizaciones, instituciones, estados o incluso países.

En otras ocasiones, representan contenido como páginas web, palabras clave o videos. Incluso pueden representar ubicaciones o eventos físicos o virtuales. Los vértices a menudo se corresponden con los bloques de construcción primarios de las plataformas de redes sociales como páginas en wikis, amigos en sitios de redes sociales y publicaciones o autores en blogs.

Edges

Los bordes, también conocidos como enlaces, lazos, conexiones y relaciones, son el tejido conectivo de las redes. Una arista conecta dos vértices entre sí. Los bordes pueden representar muchos tipos diferentes de relaciones como proximidad, colaboración, parentesco, amistad, asociación comercial, citación de literatura, inversión, hipervínculo, transacción o cualquier atributo compartido (por ejemplo, personas que asistieron a la misma universidad).

Se puede decir que existe una ventaja si tiene algún estatus oficial, es reconocida por los participantes, o se observa por intercambio o interacción entre ellos. En resumen, una arista es cualquier forma de relación o conexión entre dos entidades.

Betweenness centrality

Ayuda a identificar a las personas que desempeñan un papel de "puente transversal" en una red.

La noción de rutas de conexión es fundamental para el estudio de las redes. Quizás una de las preguntas más naturales sobre dos personas en una red es "¿Qué tan lejos están?" Esta distancia se mide simplemente: la distancia entre personas que no son vecinos se mide por el menor número de saltos de vecino a vecino de uno para conectarse con el otro. Por ejemplo, las personas que no son tus vecinos, pero son vecinos de tus vecinos, están a una distancia 2 de ti, y así sucesivamente. El camino más corto entre dos personas se llama "distancia geodésica" y se utiliza en muchas métricas de centralidad. Por ejemplo, la centralidad de intermediación (Betweenness centrality) es una medida de la frecuencia con la que un vértice dado se encuentra en el camino más corto entre otros dos vértices.

Ejercicio práctico

En un ejercicio para ejemplificar cómo usar **NodeXL** para hacer netnografía y análisis cualitativo, llevamos a cabo junto con la Dra. Verónica Espinoza un análisis sobre la percepción de la obesidad en México en la red social de Twitter.

En este caso la búsqueda se realizó mediante operadores booleanos:

- obesidad OR sobrepeso OR #obesidad OR #sobrepeso OR #obesidadinfantil

Geolocalizando la búsqueda específicamente a México, de tal forma que la búsqueda se debe hacer junto con la geolocalización como aparece en la ilustración 14:

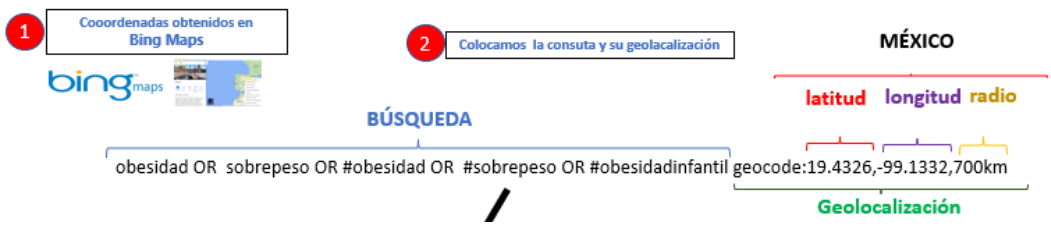


Ilustración 14. Autor: Dra. Verónica Espinoza

Tanto la búsqueda como la geolocalización se debe insertar en el importador de NodeXL:

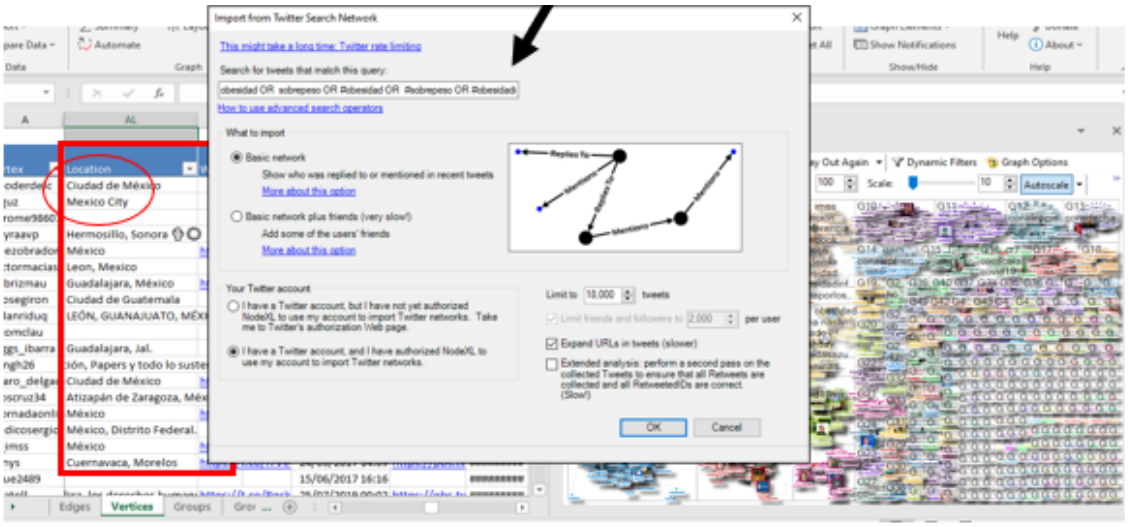


Ilustración 15. Autor: Dra. Verónica Espinoza

Una vez que hemos corrido las métricas, obtuvimos la siguiente visualización de nuestra red, en donde observamos la estructura general de la comunidad y las diferentes comunidades, para el análisis nos detuvimos en las comunidades G2 y G3.

Es recomendable usar la automatización de datos con las “recetas” disponibles en: https://www.smrfoundation.org/nodexl/automation/twitter-data-recipes/

Y utilizar este tutorial: https://www.smrfoundation.org/wp-content/uploads/2019/02/NodeXL-Pro-Tutorial-Social-network-and-content-analysis-with-Twitter-network-data-%E2%80%93-step-by-step.pdf

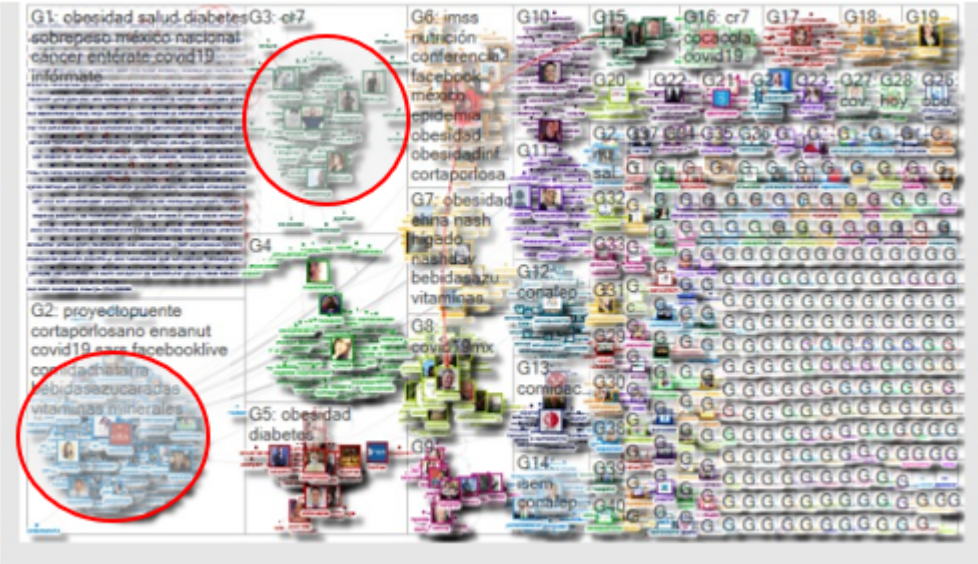


Ilustración 16. Autor: Dra. Verónica Espinoza

Ya que hemos localizado las comunidades es importante tener una ruta metodológica, el primer paso es determinar cuáles son los principales influenciadores, para lo cual debemos usar una métrica llamada **Betweenness centrality.**

Para localizar la columna de **Betweenness centrality,** es necesario primero posicionarse en la pestaña de **vértices** y luego buscar la columna.

Una vez localizada la columna **Betweenness centrality**, se deben ordenar las filas conforme a los parámetros de mayor a menor, de tal forma que veremos cómo se reordena la información en la tabla y aparecerán en primer lugar los actores sociales que visualmente se pueden identificar al centro de las comunidades, como se observa en la ilustración 17.

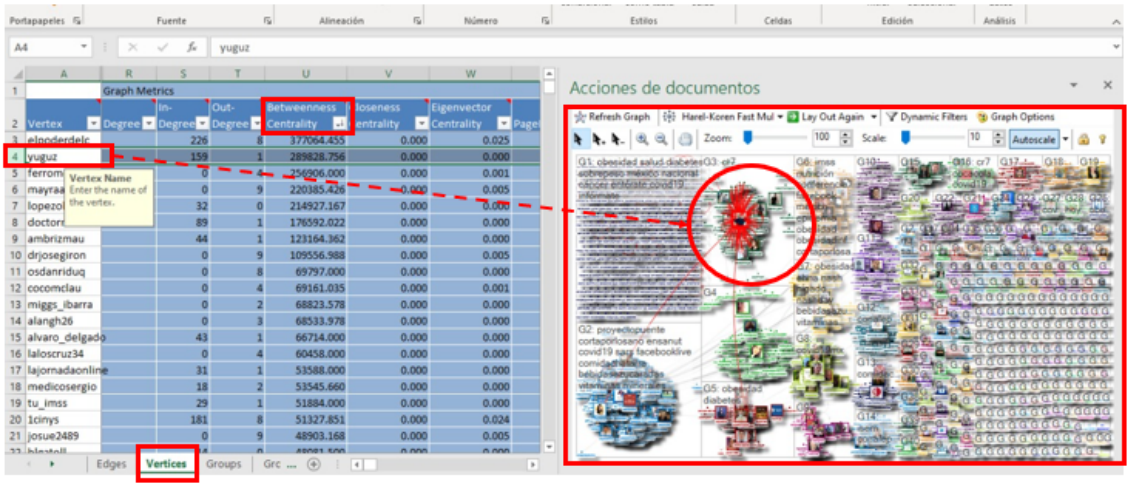


Ilustración 17: Autor Marcos de Colsa

Localizar los influenciadores es crucial para analizar qué información nos proporciona la red y cuáles son las conversaciones relevantes. Los influenciadores pueden ser de diferente naturaleza, como actores sociales pueden ser personas, organizaciones, instituciones, empresas, etc., Es importante notar que los influenciadores en este contexto son actores sociales que detonaron cierto tipo de actividad relevante en la comunidad, y su influencia es netamente contextual, es decir, está limitada por un corte de información específica de tiempo y espacio.

Los cortes que nos permite hacer Twitter son de ocho días, por lo que es recomendable hacer diferentes cortes en diferentes momentos, para ver cómo evolucionan las comunidades y los influenciadores.

Debido a que las redes y comunidades digitales son sumamente flexibles y dinámicas es muy común que los influenciadores cambien en cada corte que se haga, sin embargo, dependiendo del tema y la problemática que estemos analizando, puede haber influenciadoras que tengan una presencia constante en nuestras redes.

La primera ruta para hacer análisis cualitativo es, ya localizados los influenciadores, revisar la publicación que generó el flujo de actividad en su perfil en Twitter, en este caso vemos (como se muestra en la ilustración 17) que el Grupo 2 (G2) y el Grupo 3

(G3) conforman comunidades bien definidas, y en el centro se encuentran los influenciadores que ya identificamos mediante la métrica de **Betweenness centrality.**

Para poder revisar la publicación específica que ha activado las conversaciones en Twitter, es necesario ubicar la columna “Tweeter Page For Tweet” en la pestaña de “Edges” y dar clic en el vínculo.

En este caso la liga nos llevó a un Tweet que fue el detonador de las conversaciones y la actividad del G3 que hizo del actor social ubicado por las métricas, ser un actor relevante, es decir, influenciador, en ese corte particular sobre el tema de la obesidad en México.

El Tweet que detonó la actividad es:

“La obesidad infantil debería ser considerada maltrato”

En torno al cual se desarrolló un hilo de conversación importante, así como actividad de “Retweets” y “Me gusta”.

El Tweet por sí mismo es provocativo y eso explica que haya generado atención y actividad pública, algunos manifestándose en favor del comentario y otros en contra, es decir, desde la perspectiva del análisis cualitativo ya tenemos una primera clasificación de información, pero lo más relevante es que en este caso particular, tenemos un hilo de conversación bastante robusto que nos permite tener suficiente material de trabajo, equivalente a haber hecho uno o dos grupos de discusión.

En este caso, ya que tenemos localizado el Tweet que generó la actividad, podemos hacer copy/paste de los diferentes comentarios relevantes (dejando de lado todos aquellos comentarios “basura”) en el hilo de conversación y clasificarlos en una tabla de Excel de acuerdo con los criterios que generalmente seguimos en nuestros análisis, aquí cada investigador seguirá su ruta con la que ya está familiarizado.

Para sacar el máximo provecho de la información que hemos descargado con nuestra herramienta, nos debemos posicionar en la pestaña “Edges” y localizar la columna “Tweet”, el siguiente paso es seleccionar toda la columna, copiar y pegar en

una nueva tabla de Excel, y seleccionar los diferentes tweets de acuerdo con su contenido.

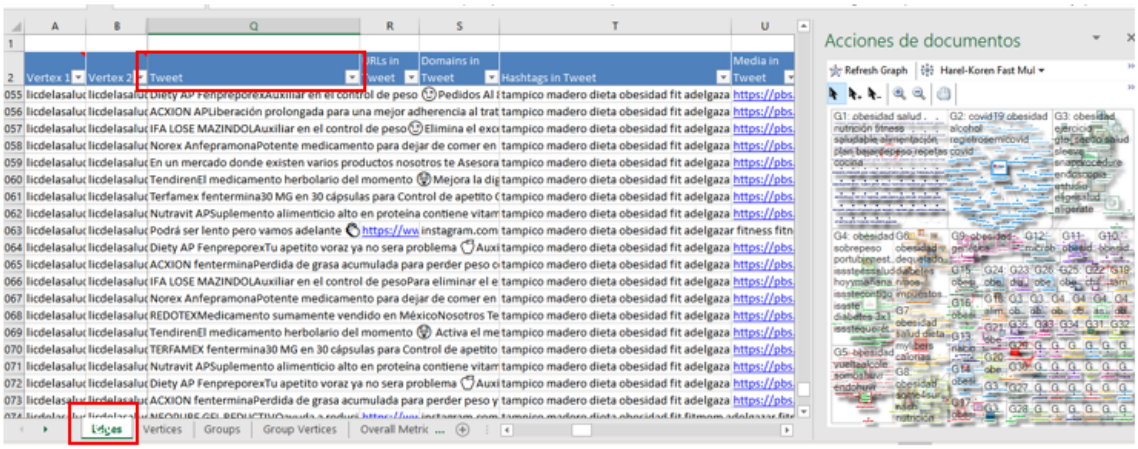


Ilustración 18: Autor Marcos De Colsa

Esta parte de “limpieza” de información es de hecho uno de los pasos más importantes ya que tenemos que separar la información basura de la información relevante, así es que, aunque nos tome un par de horas sí es importante hacer la limpieza de datos.

Cuando hemos localizado influenciadores con actividad que ha generado hilos de conversación robustos, lo que tenemos es básicamente la materia prima para nuestro análisis cualitativo.

Es importante considerar que no necesariamente cuando hay un influenciador y actividad relevante vamos a encontrar suficiente material en formato de texto, en estos casos también podemos utilizar otros recursos para nuestro análisis cualitativo como el uso de imágenes.

La ventaja de utilizar NodeXL PRO es que bajamos diferentes formatos de datos, lo cual es una de las características esenciales de la netnografía, trabajar con diferentes tipos de datos (texto, imagen, audio, video, etc.)

Si queremos hacer un análisis cualitativo a partir de las imágenes, debemos buscar en la misma pestaña de “Edges” la columna “Media in Tweet”, hacer clic en los

diferentes enlaces que ahí encontraremos y seleccionar las imágenes que nos interesan para nuestro análisis cualitativo.

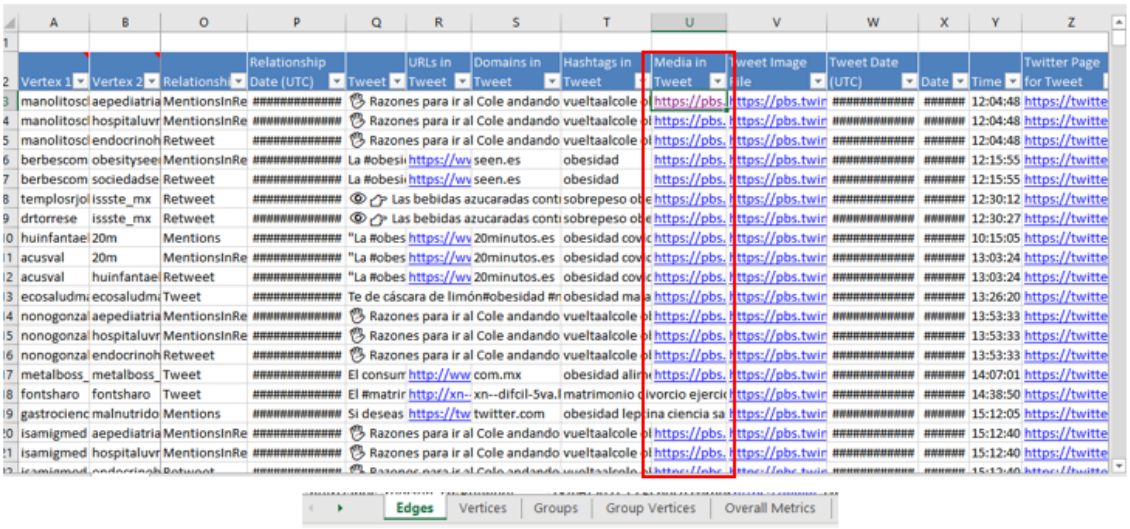


Ilustración 19: Autor Marcos De Colsa

Cuando hacemos análisis cualitativo generalmente pensamos en texto, debido a que las metodologías de investigación como las entrevistas o los grupos focales son las técnicas más extendidas, sobre todo en el medio del marketing y las consultorías de investigación de mercado, las cuales, a partir de lo expresado oralmente por los informantes, se transcriben los testimonios y lo que tenemos es una gran cantidad de texto.

Sin embargo, las imágenes complementan de forma muy poderosa el análisis textual y en ocasiones, si tenemos suficiente material, las imágenes pueden ser la principal materia de trabajo ya que las imágenes contienen referencias simbólicas complejas y muy poderosas para el análisis cualitativo.

Como lo muestra la ilustración 20, las imágenes tienen un poder evocativo sumamente profundo para analizar una infinidad de temas y problemáticas sociales, en este caso vemos una pequeña fracción de las imágenes que pudimos recolectar para analizar la problemática de la obesidad en México a partir de las representaciones simbólicas y culturales que reflejan las diferentes imágenes.

Ilustración 20: Autor Marcos De Colsa

En síntesis, hemos realizado un breve recorrido que va desde la Revolución Digital, pasando por aspectos básicos del Análisis de Redes Sociales y Netnografía, lo cual nos pone en contexto histórico y conceptual sobre la necesidad de hacer análisis cualitativo con herramientas como NodeXL que se especializa en el tratamiento de datos de redes y la cual nos permite hacer puentes metodológicos para navegar entre el mundo de los datos que en nuestro caso es la huella digital específicamente en medios sociales como Twitter y el análisis cualitativo de los datos, conversaciones, significados y representaciones culturales de básicamente cualquier tema que nos pueda interesar y que tenga expresión pública en la esfera de los medios sociales digitales.